Долгодворова Мария

Data Scientist

+79160190890

• Москва, Россия

https://github.com/knapweedss

ОПЫТ РАБОТЫ

Mar 2023 - настоящее

Разработчик чат-ботов

время

MTS AI Москва, Россия

Занимаюсь написанием јѕ и јаіср-скриптов и внедрением МЛ-моделей в чат-боты.

За время работы приняла участие в крупном проекте по тематическому

моделировании запросов, что позволило успешно подключить распознавание

наиболее значимых интентов в чат-боты и повысить GMV.

Стек: JavaScript, Python, JAICP, Gitlab, ClearML, Kibana

Nov 2021 - Feb 2023 Удаленно

Аналитик данных

Sponsr

Участвовала в создании инструментов аналитики для компании с нуля: разработала интерактивный дэшборд на flask, подготовила витрины PostgreSQL и настроила миграцию данных для дэшборда в Looker Studio. Кроме того, занималась исследованиями аудитории: сентимент-анализ комментариев, анализ логов

оплаты, исследование вовлеченности пользователей.

Стек: Python, Gitlab, PostgreSQL, Clickhouse, MySQL, Kibana

НАВЫКИ

Python

PyTorch, TensorFlow, scikit-learn, xgboost, catboost, lightgbm, nltk, spacy, natasha, gensim, plotly, matplotlib, seaborn, flask, dash, streamlit, pandas, numpy

PostgreSQL, ClickHouse, MySQL, MongoDB, Oracle Database

Git

JavaScript

Docker

R

ОБРАЗОВАНИЕ

2023 - 2025 Магистратура «Анализ Данных в Биологии и Медицине»

Москва, Россия Национальный исследовательский университет «Высшая школа экономики»

2018 - 2023 Бакалавриат «Фундаментальная и Компьютерная Лингвистика» Москва, Россия Национальный исследовательский университет «Высшая школа экономики»

ЗНАНИЕ ЯЗЫКОВ

Английский, С1

Французский, А2

ПРОЕКТЫ

Meadow Mari Spoken Corpus

Anna Volkova, Aigul Zakirova, Mikhail Voronov, Maria Dolgodvorova, Zinaida Klyucheva, Svetlana Kokoreva, Ilya Makarchuk, Irina Khomchenkova, Timofey Arkhangelskiy, Elena Sokur. Spoken corpus of Meadow Mari (as spoken in the village of Staryj Toryal, Novyj Torjal district, Mari El Republic, Russia). Moscow: Linguistic Convergence Laboratory, HSE University.