



Марченко Андрей
Junior Data Scientist
21 год

+7-967-123-25-25
marcheanin1576@gmail.com
GitHub Profile

ОБРАЗОВАНИЕ

- МГТУ им. Н. Э. Баумана** 2021-2025
ИУ9 Теоретическая информатика и компьютерные технологии (ПМИ) Москва
ВКР: Разработка алгоритма автоматического формирования модели кредитного скоринга
- VK Education – Отделение при МГТУ им. Н.Э. Баумана** 2023-2024
Программа: Машинное обучение. Основные курсы: Базовый ML, Нейронные Сети, DWH

ПРОЕКТЫ

- AutoML для кредитного скоринга** Дипломная работа, Весна 2025 [Source Code](#)
 - Система автоматического формирования полного пайплайна для задачи бинарной/многоклассовой классификации.
 - Оптимизация на основе генетического алгоритма, кодирующего 22-генную хромосому пайплайна.
 - Включает 5 вариативных этапов предобработки данных и 4 модели машинного обучения.
 - Реализован удобный интерфейс CLI и сериализация пайплайна для дальнейшего использования.
 - **Используемые технологии:** Python, Scikit-learn, Pandas, Numpy, Imbalanced-learn, Gensim, Tensorflow
- Анализ мемов на hateful содержание** Pet-проект, Декабрь 2023 [Source Code](#)
 - Сервис, анализирующий мемы (картинка+текст) на ненавистническое содержание.
 - Разработал объединение свёрточной и текстовой моделей с помощью классификационной головы.
 - **Используемые технологии:** Python, Pandas, PyTorch, Lightning
- ClusterText** Pet-проект, Весна 2024 [Source Code](#)
 - Web-приложение для кластеризации текстов, выделения ключевых слов и расстановки тегов.
 - Разработал модель для кластеризации текстов, относящую текст к одному или нескольким из 15 кластеров.
 - Модель состоит из эмбеддера на основе SBert и KMeans для обучения на 60к статьях.
 - **Используемые технологии:** Python, Pandas, Clustergram, KMeans, Transformers, PyTorch
- Многокритериальная оптимизация в задаче кредитного скоринга** Курсовой проект, Декабрь 2024 [Source Code](#)
 - Построение многокритериальной модели для оптимизации разноплановых loss-функций.
 - Применение генетических алгоритмов для обработки гетерогенных данных и разработка собственного алгоритма оптимизации.
 - **Используемые технологии:** Python, Pandas, Tensorflow, PyTorch

Опыт

- Компания ИНВЕРСИЯ** Июль 2022 – Февраль 2023 Москва
Программист, Java, PL/SQL
 - Работал в связке Java – SQL, решал задачи как во фронтенд, так и в бэкенд частях
- Газпромбанк** Июнь 2024 – наст. время Москва
Инженер разработки, Java
 - Микросервисная разработка на языке Java.
 - Разработка сложных структур баз данных PostgreSQL.

ТЕХНИЧЕСКИЕ НАВЫКИ

Языки программирования: Python (продвинутый уровень, опыт в разработке production-решений), SQL (написание сложных запросов с использованием оконных функций и CTE), C++ (опыт спортивного программирования и разработки проектов), Java (опыт коммерческой разработки).

Библиотеки и фреймворки: PyTorch, Scikit-learn, Pandas, Numpy, Matplotlib, Seaborn, Transformers, Lightning, Imbalanced-learn, Gensim, TensorFlow, FastAPI.

Инструменты: Git, Docker, VS Code, PyCharm

Навыки: Машинное обучение, глубокое обучение, NLP, компьютерное зрение, A/B тестирование, статистический анализ, алгоритмы и структуры данных.

Языки: Русский (родной), Английский (C1), Немецкий (A2)

СИЛЬНЫЕ СТОРОНЫ И ИНТЕРЕСЫ

Сильные стороны: усердие, исполнительность, ответственность, трудолюбие, способность к обучению и развитию, чувство юмора

Увлечения: настольный теннис, шахматы (I взрослый разряд), футбол, робототехника.