

МАРАТ ГАСАНОВ

ML-Разработчик

✉ maratgasanov6@gmail.com ☎ +7 989 881-13-36 🌐 github.com/gassannov 📧 t.me/ggassannovv

ОБО МНЕ

AI/ML-разработчик. В основном работал с генеративными моделями для текста и изображений. Дообучал модели под различные задачи, включая ответы на вопросы, суммаризацию и перевод текстов. Также занимался дообучением image generation моделей для генерации изображений в заданных стилях. Строил пайплайны генерации фонов в ComfyUI и выводил их в production. Оптимизировал инференс моделей в Triton Inference Server.

ОПЫТ РАБОТЫ

redrocket.ai, ML-РАЗРАБОТЧИК 01/2023 – 08/2023

- Разрабатывал и внедрял модель для генерации картинок в production
- Настраивал RabbitMq для работы с моделью и писал балансировку нагрузки Python · StableDiffusion · RabbitMq · VastAPI

ПАО “СберБанк”, DATA-ANALYTIC 08/2023 – 11/2023

- ML для аудита
- Поиск различных противоречий и неточностей с помощью ML-алгоритмов, работа с окр для оцифровки документов, поиск аномалий, задача прогнозирования спроса
- Spark, Hadoop, Hive, GreenPlum, Pandas, Python

Студия Артемия Лебедева, ML-РАЗРАБОТЧИК 11/2023 – 11/2024

- Разработка и поддержка моделей для генерации логотипов
- MLOps - задачи по распределению нагрузки и поддержки работоспособности
- Дообучение моделей под генерацию картинок в определенных стилях
- Работа с LLM моделями, для интерпретации пользовательских запросов

Ozon, ML-РАЗРАБОТЧИК 11/2024 – Present

- Дообучение моделей машинного перевода для различных доменов, включая UGC-чаты и описания товаров. Достигнуто качество, превышающее результаты Яндекс Переводчика на целевых данных.
- Оптимизация инференса LLM в связке Triton Inference Server + vLLM. Увеличил выдерживаемую нагрузку по RPS в 20 раз.
- Разработка пайплайнов генерации фонов и объектов по аналогии, а также генерации фонов на основе текстовых промптов.
- Работа с LLM-моделями для автоматической генерации промптов для создания фонов.
- Вывод ComfyUI-пайплайнов в production-среду.

ОБРАЗОВАНИЕ

МГТУ ИМ. Н.Э. БАУМАНА 2020 – 2024

VK образование в МГТУ им. Н.Э. БАУМАНА. 01/2022 – 06/2023

НАВЫКИ

[Triton Inference Server](#) | [vLLM](#) | [LLM models](#) | [ComfyUI](#) | [StableDiffusion](#) |

[Docker](#) · [Docker Compose](#) | [SQL](#) · [MySQL](#) · [PostgreSQL](#)