Антон Е. (Anton E.)

Прикладной ML-инженер | Модели и синтетические данные

Россия | Открыт к удалённой работе и релокации

unstonio@gmail.com | unston.io/ru

github.com/unstonio | huggingface.co/unstonio

О себе

Независимый ML-инженер с опытом практической работы с ИИ и ML более шести лет: самостоятельные исследования, фриланс и коммерческие проекты. Обучал и интегрировал кастомные модели для американского стартапа Smartist. Готовлю синтетические данные, обучаю и оцениваю специализированные модели, упаковываю инференс для локальных приложений и браузера.

Технические навыки

Модели PyTorch, Hugging Face Transformers, Sentence Transformers, авторегрессионные модели, дообучение

Данные и оценка Синтетические данные, отбор, эмбеддинги подписей, фильтрация, воспроизводимые сплиты, отложенная валидация, разбор генераций

Инференс FastAPI, ONNX, ONNX Runtime Web, INT8-квантизация, KV-кэш, локальный запуск

Программирование Python, NumPy, JavaScript, Git

Избранные технические проекты

PixelGPT 24×24

Самостоятельное исследование | Опубликованы модель, датасеты и демо

- Разработал авторегрессионный трансформер на 68 млн параметров для генерации 24×24 по тексту и пятицветной палитре. Описанный запуск обучения использовал около 218 тыс. примеров.
- Обучал в PyTorch с кросс-энтропией следующего токена и валидацией на отложенной выборке. Описанный запуск на RTX 5090 занял около 30 минут. Проверял генерации с разными запросами и палитрами.
- Опубликовал 20 тыс. примеров по CC BY 4.0 и отдельный коммерческий релиз из 216 838 изображений с подписями, метаданными, сплитами и офлайн-просмотрщиком с поиском.
- Выпустил локальное приложение на FastAPI с потоковой генерацией и редактором. Перенёс инференс в Web Worker с ONNX Runtime Web и INT8-декодером.

Кейс | GitHub | Демо | Открытая выборка | Полный датасет

PixelGPT Zero

Исследовательский эксперимент | Доступно демо

Обучил многоуровневую модель MicroVAR на процедурных и отобранных данных. Экспортировал INT8 ONNX (1,6 МБ) для генерации в браузере с KV-кэшем.

Демо в браузере

Профессиональный опыт

Smartist | AI/ML-инженер

Стартап из США | Удалённый контрактный проект

- Готовил датасеты, проводил эксперименты, обучал и дообучал кастомные модели для приложения Smartist.
- Упаковывал инференс и интегрировал модели в продуктовый пайплайн приложения.
- Проверял результаты в приложении и дорабатывал модели с учётом требований продукта и обратной связи.

Независимый ML-инженер | Фриланс-разработчик

Самостоятельные исследования с 2020 года | Фриланс с 2023 года

- Разрабатываю специализированные генеративные модели и пайплайны синтетических данных, от программной генерации и отбора до валидации и инференса.
- Публикую воспроизводимые демо, код моделей и датасеты. Для клиентов создаю интеграции с агентами, поиском и API, проверяю работу и готовлю документацию.
- Разработал Slava+, интерактивный дизайн-инструмент с трекингом рук в браузере и движком геометрической композиции.

Демо Slava+

Студия Артемия Лебедева | Дизайн и креативные технологии

Москва | Предыдущая роль

- Готовил визуальные материалы, цифровые ассеты и презентации для студийного производства. Проверял качество и согласованность итоговых материалов.
- Работал в студийной команде над визуальными системами и графикой для фирменного стиля.

Другие исследования

Tiny Vector

Исследовательский прототип | Демо в браузере

Обучил два декодера по 2,24 млн параметров для сравнения отрезков и квадратичных кривых на процедурном датасете иконок. Экспортировал модели в ONNX для генерации редактируемых SVG в браузере с декодированием по грамматике.

Интерактивный прототип

Образование

Самостоятельное обучение машинному обучению и разработке ПО на основе учебных материалов, экспериментов и прикладных проектов.

Языки

Русский (родной). Английский (чтение и написание технической документации).