



Evdokimov Alexander

Мужчина, 35 лет, родился 7 июля 1989

+7 (916) 6777147

snakextntx@gmail.com — предпочитаемый способ связи

Другой сайт: <https://github.com/snakerzr>

Проживает: Москва

Гражданство: Россия, есть разрешение на работу: Россия

Готов к переезду, готов к командировкам

Желаемая должность и зарплата

Data scientist (NLP/LLM)

Специализации:

- Дата-сайентист
- Другое

Занятость: полная занятость, частичная занятость, проектная работа

График работы: полный день, гибкий график, удаленная работа

Желательное время в пути до работы: не имеет значения

Опыт работы — 3 года 4 месяца

Март 2024 —
настоящее время
9 месяцев

Эвотор

Москва

Информационные технологии, системная интеграция, интернет

- Разработка программного обеспечения
- Системная интеграция, автоматизации технологических и бизнес-процессов предприятия, ИТ-консалтинг

Ведущий разработчик-лингвист

- Участвовал в разработке AI-помощника на базе современных LLM и векторных баз данных
- После критического анализа современных методик автоматической оценки предложил и ввел в эксплуатацию метрики для анализа качества ответов помощника, отдельно для ретриверной и генеративной частей
- Исследовал и внедрял передовые методы улучшения поисковой и генеративной частей RAG для повышения точности и релевантности ответов

Октябрь 2023 —
Март 2024
6 месяцев

ООО КлаудФэктори

Москва

Информационные технологии, системная интеграция, интернет

- Разработка программного обеспечения
- Системная интеграция, автоматизации технологических и бизнес-процессов предприятия, ИТ-консалтинг

Data Scientist

- Исследовал применимость современных методов оптимизации инвестиционного портфеля из научных статей, основанных на эволюционных алгоритмах и машинном обучении
- Разрабатывал прототипы для распознавания структурированных данных в PDF-документах, Excel-файлах и изображениях с целью унифицированной консолидации данных для финансового анализа

Август 2021 —
Сентябрь 2023
2 года 2 месяца

Индивидуальное предпринимательство / частная практика / фриланс

Data Scientist (Self-taught / Educational Projects / Hackathons)

Разрабатывал и выполнял несколько учебных проектов в области анализа данных и машинного обучения, демонстрируя свои навыки в областях:

- работы с табличными данными,
- временными рядами,
- с текстом, естественным языком и большими языковыми моделями (prompt engineering, ChatGPT, YandexGPT),
- изображениями,
- DevOps.

Стек технологий: LangChain, YandexGPT, OpenAI, prompt engineering, Pandas, Numpy, Seaborn, Matplotlib, Sklearn, Scipy, Optuna, Hyperopt, Catboost, Lightgbm, XGBoost, SHAP, Pytorch, PySpark, Docker, Clickhouse, RabbitMQ, streamlit, PyPortfolioOpt.

=====

- Хакатон: Цифровой прорыв - 3 место в лиге новичков- Разработка помощника написания инвестиционной презентации для стартапов

Проблема: Зачастую стартапы представляют из себя молодые проекты с небольшой командой, у которой не всегда достаточно компетенций или денежных средств для полноценной и качественной проработки данного документа. Поэтому необходимо создать сервис, который поможет составить презентацию, провести необходимый финансовый анализ и визуальное оформление.

Решение: Создан сервис на основе gpt 3,5 и 4, в котором пользователь описывает свой стартап, а сервис проводит дополнительные финансовые исследования и генерирует готовую презентацию, автоматически структурируя информацию на слайды.

=====

- Хакатон: AI Generative Product Hackathon - Разработка чатбота для Яндекс Лавки на основе YandexGPT и LangChain

Проблема: Разработать решение на основе ИИ, которое поможет давать персонализированные ответы и оптимизирует процессы службы поддержки сервиса, дав возможность специалистам сфокусироваться на более сложных задачах.

Решение: Создана цепочка взаимодействия чатбота на основе YandexGPT с клиентом, которая способна обработать жалобы на проблемы с доставкой или товарами с учетом настроения клиента. Запланирована реализация проверки сообщений ботов на наличие опасного содержания, соответствию tone of voice Лавки, перехвата инструкций. Проведена оценка возможности масштабирования и нагрузки на сервис в часы пик.

=====

- Хакатон: Marking Hack - 2-е место - сервис прогноза продаж и детекции аномалий

Проблема: Необходимо на основе истории о вводе, транспортировке и выводе товаров из системы честный знак создать сервис, который приносил бы пользу государству и/или бизнесу. Оценивалась работоспособность сервиса, интерфейс, проработка концепции и пользы, качество питча.

Решение: Был создан сервис по прогнозированию продаж конкретного товара, в конкретном субъекте федерации на несколько месяцев вперед. И сервис по детекции аномалий в этих данных, которые могли бы свидетельствовать об уходе от налогов/мошеннических действиях с помощью магии и волшебства.

=====

- Инфраструктура для сбора данных с биржи [RabbitMQ, Clickhouse, docker]

Проблема: Для отработки торговой гипотезы мне необходима историческая информация о лимитных ордерах, который мой брокер не представляет. Только текущее состояние. Задача - создать докер контейнеры, которые бы запрашивали и сохраняли информацию о лимитных ордерах.

Решение: Были созданы 2 образа: образ со сборщиком информации с помощью API Финама, образ с обработчиком и сохранением этой информации в clickhouse. В качестве брокера сообщений используется RabbitMQ. Также созданы дашборды для контроля качества поступающих данных.

=====

Работа с табличными данными и временными рядами:

- Предсказание оттока клиентов телеком компании
- Анализ временных рядов, прогнозирование заказов такси
- Прогнозирование стоимости автомобиля по характеристикам
- Прогнозирование стоимости жилья в PySpark
- Хакатон: Первой Грузовой компании - прогнозирование длительности нахождения вагона в пути
- Хакатон: ВкусВилл - исследование спроса с помощью SHAP

=====

Работа с текстом:

- Определение токсичных комментариев [toxic-bert, pytorch]
- Хакатон Росэлторга: поиск товаров, услуг, прошедших тендеров и аналитика для поставщиков и заказчиков

=====

Работа с изображениями:

- Поиск изображения по описанию [OpenAI CLIP]

=====

Дополнительно (финансы):

- Maximum value extraction from CPAMM [crypto, оптимизация]
- Template for Market value-weighted index [fintech, pandas]
- Оптимизация инвестиционного портфеля [PyPortfolioOpt, streamlit]

Образование

Высшее

2012

Национальный институт бизнеса, Москва

Факультет экономики и управления, Менеджмент организации

Повышение квалификации, курсы

2022

Специалист по Data Science плюс

АНО ДПО «Образовательные технологии Яндекса», Data Science

Электронные сертификаты

2021

EF SET English Certificate 76/100 (C2 Proficient)

Навыки

Знание языков

Русский — Родной

Английский — C2 — В совершенстве

Навыки

Python

Git

Docker

PyTorch

ML

NLP

LLM

RAG

Дополнительная информация

Обо мне

tg: <https://t.me/snakerzr>

github: <https://github.com/snakerzr>

- Обожаю когнитивные науки и философию сознания
- Нацелен на реальный результат, а не на создание технологии ради технологии
- Стремлюсь к развитию в управленческой роли
- Серийный хакатонщик с призовыми местами
- Состою в закрытом сообществе Data Science: участвуем в хакатонах, проводим образовательные лекции и делимся опытом
- Участвую и помогаю организовывать локальные образовательные IT-события, проводил тематические лекции по Data Science
- В прошлом большой опыт любительских театральных выступлений