

Данил Кизеев

Санкт-Петербург, Россия

☎ +79040283459 • ✉ kizeevdaniil@yandex.ru

Технологии

python, оптимизация, данные, машинное обучение, computer vision, NLP, статистика, numpy, pandas, PostgreSQL, ClickHouse, scipy, transformers, pytorch, C++, Rust, Docker.

Опыт

ЦИТМ Экспонента + Indeera

Инженер ИИ

Москва, Россия

2023 Окт – 2025 Фев

Создал: API по агрегации и обработке данных, поиск оптимальной цены для продавца, поиск конкурентных товаров, выделение трендов и сезонностей, фреймворк разметки.

- Сбор данных
 - Несколько источников: собственный + wb + mpstats;
 - Исправление ошибок в процессе сбора; дополнение данных для гладкого временного ряда; проводились необходимые агрегации в зависимости от места торговли;
 - Итог — создание разных комбинаций источников под разные нужды.
- Поиск конкурентов
 - Товар переводится в эмбединг по изображению, названию и описанию;
 - Основа на самообученном CLIP на CC12M+COCO переведённые с помощью Helsinki-NLP, внутри которого ResNet и DistilBERT;
 - В качестве развития сделана сиамская сеть с агрегацией трешхолдов по категории.
 - Итог — внедрение в интерфейс SKU потенциальных конкурентов.
- Поиск оптимальной цены (ДЦО)
 - Решение с помощью бандитов (MAB + CMAB). Модифицированные UCB и Thompson sampling;
 - Дисконтирование заказов с учётом тренда + сезонности. НС — BiLSTM + Transformers;
 - Обучающая выборка — данные продаж + своя модель рынка;
 - Модель рынка на распределении Пуассона, линейным смещением для тренда и двух гармониках Фурье для сезонности.
- Docker + Flask + Gunicorn + Я.Облако для тестов и прода. Выкатка в интерфейс заказчика.

Точка сборки

Геймдев-инженер

Пермь, Россия

2022 Сен – 2023 Мар

Создание матчмейкера с помощью методов оптимизации — генетический алгоритм и дифференциальная эволюция, кастомные функции кроссинговера и селекции для эвристических метрик (люди с одинаковыми интересами, одна девушка в команде); создание пайплайнов, визуала и API для сбора данных; визуализации на egi и masroquad; язык Rust; создание игрового движка.

Нурр

Стажер Data Science

Санкт-Петербург

2021 Апр – 2021 Июнь

Автоматическая категоризация постов соц. сети для рекомендательной системы. По тексту (UMAP для сжатия) и изображению из поста получается эмбединг (word2vec+tf-idf/BERT + BERTopic), кластеризацией (HDBSCAN, kmeans) с последующим семплингом из категории нового поста определяется похожий и рекомендуется. На второй итерации произведено улучшение (CLIP на ResNet + DistilBERT) для эмбединговой части.

Образование

Университет.....

Университет ИТМО, Россия

Аналитика данных, Магистратура

2023 – 2025

Санкт-Петербургский Государственный Университет, СПбГУ, Россия

Прикладная математика и информатика, Бакалавриат

2018 – 2021

Тверской Государственный Университет, ТвГУ, Россия

Компьютерная безопасность, Бакалавриат

2017 – 2018

Проекты

Русско-карельский переводчик: [Github](#)

Создал первый онлайн-переводчик для ливвиковского наречия на основе NLLB, mBART и Qwen3. Дополнительно сделал телеграм-бота для разметки данных на параллельный корпус. Сбор данных с помощью собственных парсеров сайта, словаря и википедии. На собранном корпусе в 50к получилось побить существующее решение.

Mindmaps for everything: [Github](#)

Создал майндмапы для лучшего понимания и запоминания предметов (вдохновлялся «Цеттелькастеном»). В репозитории есть карты для статистики, NLP, оптимизации и глубокого обучения, доступны на русском и английском языках.

Рекомендер музыки на основании слов: [Github](#)

Using doc2vec and BERT embeddings for recommendation based on songs lyrics.

Создание эмбедингов на основании doc2vec и DistilBERT для последующих рекомендаций. Визуализация с помощью UMAP и Plotly. Создал телеграм-бота в качестве оболочки.

Контакты.....

GitHub: <https://github.com/breadfan>

StackOverflow: <https://stackoverflow.com/users/9850300/taciturno>

LinkedIn: <https://www.linkedin.com/in/rocauc>

LeetCode: <https://leetcode.com/breadfan/>

Telegram: <https://t.me/rocauc>