

КАЛИНИН НИКИТА

CONTACT

+7 993 287 07 68
Kalininskiy@yandex.ru
<https://github.com/NikitaKQ>
@kal1n1n_n

HARD SKILLS

Python
Data science tools
(pandas, numpy, plotly,
matplotlib, seaborn, scikit-learn)
SQL
C++
R
Computer vision
NLP, RNN, LLM, Hugging Face
English - C1
Алгоритмы и структуры данных
Microsoft office
A/B тестирование
Управление командой
Участие в кейсчемпионатах:

- Cup Moscow 2023 (финал)
- Cup IT 2024 (финал)

Участие в хакатонах

- Хакатон от компании "Самолет" 2023
- VK x HSE Data Hack

SOFT SKILLS

Управление командой
Работа в команде над совместными проектами
Организация мероприятий
Менеджмент
Волонтерство в университете
Был ассистентом по C++ в Зимней Олимпиадной школе от МФТИ

ОБРАЗОВАНИЕ

Бакалавриат

НИУ ВШЭ - Москва

2022-2024

Факультет компьютерных наук

Прикладной анализ данных

Пройденные курсы:

- JetBrains Academy SQL and Databases
- Deep learning(CV) - Stepik МФТИ
- Deep learning(LM) - Stepik МФТИ

ПРОЕКТЫ

ПРИЛОЖЕНИЕ НА QT ДЛЯ АНАЛИЗА И РАБОТЫ С ДАННЫМИ

Технологии: C++, библиотека QT, Qtcreator

2023

Сделал приложение, в котором можно работать с данными: добавлять убирать строки и столбцы датасета, изменять данные в строках, фильтровать, сортировать разные типы данных. Приложение поддерживало несколько языков, а также имело свой логотип.

СТАЖЕР БИЗНЕС АНАЛИТИК

GuruLeads

2023

Технологии: Jira, Figma, YandexWiki, Powerpoint

Проводил анализ сайта компании, проверял интуитивность работы интерфейса. Придумывал концепцию работы раздела поддержки. Ставил задачи дизайнеру и программистам в Jira. Делал анализ конкурентов. Строил воронку продаж. Составлял A/B тестирования.

ГРУППОВОЙ ПРОЕКТ ПО COMPUTER VISION

2023 - 2024

Технологии: Pytorch, Optuna, torchvision, cuda

В проекте занимался написание модели для идентификации дипфейков, настраивал генерацию модели и гиперпараметров с помощью алгоритмов optuna. Занимался реализацией архитектуры генеративной модели GAN.

АНАЛИЗ ДАТАСЕТА С ДАННЫМИ О ТРЕКАХ В SPOTIFY

2022

Технологии: Numpy, Plotly, Pandas, Seaborn и Matplotlib

Составлял гипотезы о популярности музыки в зависимости от звучания тембра и других характеристик звука. Проводил очистку, структуризацию данных и строил графики для графического доказательства гипотез. В конце выступал с защитой проекта.

СОЗДАНИЕ КОНЦЕПЦИИ ИДЕАЛЬНОЙ ПЛАНОГРАММЫ ДЛЯ P&G CUP IT ОТБОРОЧНЫЕ

2024

Технологии: Numpy, Plotly, Pandas, Seaborn и Matplotlib, MySQL, Miro

Составлял план работы команды, распределял задачи между членами команды, занимался общим планированием проекта. Проводил анализ датасетов о кругообороте продуктов от склада до продажи в на кассе. Составлял AB тестирования и ANOVA тесты для выборки, придумывал алгоритм работы идеальной планогаммы.

ML ДЛЯ ВЫЯВЛЕНИЕ НЕКАЧЕСТВЕННЫХ ОТЧЕТОВ ОТ ДЗО CUP IT ФИНАЛ

2024

Технологии: Numpy, Plotly, Pandas, Seaborn и Matplotlib, sklearn

Составлял план работы команды, распределял задачи между членами команды, занимался общим планированием проекта. Проводил обработку и анализ данных о поставках отчетов ДЗО Сбербанк для кейсчемпионата, строил различные модели для выявления и предсказания невалидных поставок

ХАКАТОН VK X HSE DATA HACK

2024

Технологии: Pytorch, Hugging Face, saiga LLM, RuBERT, word2vec, VK Cloud, sklearn, numpy, pandas, gensim, nltk, CatBoost

На хакатоне был капитаном команды. Занимался разработкой модели для определения класса новостной статьи (Политика, Экономика, История, Дом) и ключевых слов. Использовали предобученного на русском языке ruBERT и дообучали на размеченном датасете с данными, а также с помощью LLM saiga занимались выделением ключевых слов. Пробовал использовать word2vec и CatBoost для выделения ключевых слов.