

СЕРГЕЙ ВИКУЛИН



+7 913-707-32-86



savikulin@gmail.com



Telegram



GitHub

ЖЕЛАЕМАЯ ДОЛЖНОСТЬ: ПРОДУКТОВЫЙ АНАЛИТИК

Меня зовут Сергей, я студент 4 курса экономического факультета НИУ ВШЭ. В школе я увлекался математикой и успешно принимал участие в различных олимпиадах, став призёром олимпиады "Ломоносов" и победителем Всесибирской олимпиады школьников по математике. На 2 курсе университета я осознал, что теоретические концепты экономики не всегда ориентированы на бизнес, и мой интерес переключился на анализ данных. Я стал увлекаться А-В-тестированием, визуализацией данных и машинным обучением, прошел курсы "Основы статистики и А/В-тестирования" от Яндекс практикума, "Симулятор SQL" от Karpro Courses и "Анализ данных в индустрии" от Тинькофф. Я владею основными техническими навыками аналитика данных, такими как Python, PostgreSQL, Tableau, Datalens. Мой основной интерес сейчас сосредоточен на задачах в области онлайн-экспериментов, таких как региональные А-В-тесты, динамическое ценообразование и предиктивная классификация. В данный момент я активно развиваю свои технические навыки и ищу работу в области аналитики данных, готов применить свой опыт и знания для успешного решения бизнес-задач.

ПРОФЕССИОНАЛЬНЫЙ ОПЫТ

Kaggle (учебный опыт):

A-B-testing of sales conversion rate for online shop | [Github](#)

Апрель 2024

Стек: *python (pandas, seaborn, matplotlib), A-B-тестирование, продуктовые метрики*

Цель: Проверка разбиения и интерпретация результатов А-В-тестирования

Применимость: Проведение А-В-тестирования в онлайн экспериментах для увеличения показателей бизнеса

- Проверил разбиение тестов на ошибки и репрезентативность выборок
- Визуализировал основные числовые и категориальные переменные и описал их распределение
- Рассчитал продуктовые метрики, такие как DAU, WAU, MAU, ARPU, ARPPU, AOV, sticky factor, rolling retention, конверсия в покупку, среднее количество сессий, среднее количество действий за сессию, среднее время между сессиями, среднее количество дней «жизни» пользователя
- Провел z-тест для разницы конверсий в покупку
- Выполнил t-тест для разности ARPU и для разности ARPPU с помощью бутстрапа

Exploratory data analysis of Zomato food delivery app

 | [Github](#)

Апрель 2024

Стек: *python (pandas, scipy, seaborn, matplotlib), статистика и теория вероятностей, продуктовые метрики*

Цель: Предварительный анализ данных бизнеса

Применимость: Анализ данных бизнеса заказчика, подрядчика для выявления преимуществ и недостатков

- Сделал предварительную обработку данных (очистка данных от пропусков, создание больших когорт по типам кухонь и видам блюд, объединение таблиц, удаление выбросов по сумме и количеству продаж)
- Визуализировал признаки и описал основные закономерности их распределения
- Протестировал гипотезы о равенстве средних с помощью бакетного преобразования и линеаризации
- Сегментировал пользователей по RFM
- Построил линейную регрессию для суммы заказа и проверил значимость коэффициентов
- Посчитал ключевые продуктовые метрики

Visualization of Brazilian E-commerce shop key metrics

 | [Tableau Datalens](#)

Август 2023

Стек: *Tableau, Datalens*

Цель: Построение дашборда для отображения основных метрик бизнеса

Применимость: Построение дашбордов для руководства по KPI или для отчетности

- Вычислил основные показатели бизнеса (выручка, количество продаж, ARPPU, средний чек, средняя сумма доставки, DAU, WAU, MAU)
- Построил график динамики продаж, карту продаж, таблицу основных показателей по категориям товаров и городам покупателей, гистограммы DAU, WAU, MAU
- Создал дашборд с селекторами, позволяющими выбрать временной интервал и штат покупателя

Selection of hyperparameters for machine learning models

 | [Python](#) [Github](#)

Август 2023

Стек: *python (pandas, sklearn, matplotlib, seaborn), машинное обучение*

Цель: Поиск наилучшей модели и подходящих гиперпараметров для задачи классификации

Применимость: Применение методов машинного обучения для решения задачи классификации в бизнесе

- Обучил 5 моделей классификации (KNN, Decision Tree, SGD Linear Classifier, Random Forest, Gradient Boosting) на вещественных признаках
- Произвел масштабирование признаков и добавил категориальные переменные в модели
- Подобрал наилучшие комбинации гиперпараметров
- Нашел новые полезные объясняющие переменные с помощью создания полиномиальных признаков и их отбора посредством фильтрационных методов, жадного отбора признаков, встроенного в модель отбора признаков
- Смешал модели - Gradient Boosting и SGD Linear Classifier, показавшие наилучшее качество на кросс-валидации

ОБРАЗОВАНИЕ

Национальный исследовательский университет «Высшая школа экономики» Сентябрь 2020 – Июль 2024

Факультет экономических наук - Бакалаврская программа «Экономика»

Основные курсы: Теория вероятностей и статистика, Линейная алгебра, Математический анализ, Машинное обучение, Микроэкономика, Макроэкономика, Эконометрика, Теория и практика онлайн экспериментов

ТЕХНИЧЕСКИЕ НАВЫКИ И ИНТЕРЕСЫ

Языки программирования: Python (pandas, numpy, scipy, sklearn, matplotlib, seaborn), SQL (aggregated functions, conditions, joins, subqueries, window functions)

BI-инструменты: Tableau, DataLens

Технические навыки: Теория вероятностей и статистика, А-В-тестирование, Машинное обучение, Глубинное обучение, Визуализация данных

Языки: Английский (upper-intermediate), Русский (native)

Личные интересы: Футбол, Шахматы, Тхэквондо, Кикбоксинг, Кино, Путешествия