

Зигангиров Михаил Родионович

Data Analyst | Data Scientist

Email: zig@ngirov.ru

Telegram: @hallaren

Образование

ИТМО, Санкт-Петербург

2022 – 2026

(Факультет технологического менеджмента и инноваций, направление «Бизнес-информатика»)

Навыки

- **SQL:** оконные функции, группировка, подзапросы, оптимизация запросов
- **Python:** Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Math, scipy, TensorFlow, PyTorch
- **Алгоритмы и модели машинного обучения:** линейная регрессия, логистическая регрессия, деревья решений, случайные леса, градиентный бустинг (XGBoost, LightGBM, CatBoost)
- **Математика и статистика:** лин. алгебра, мат. анализ, вероятности, распределения, гипотезы, корреляции, комбинаторика, оптимизация.
- **A/B тестирование:** проектирование и анализ экспериментов
- **Git/GitHub**

Курсы

- Яндекс Практикум “Аналитик данных”
- ИТМО “Машинное обучение и анализ данных”

Проекты

Анализ среднего времени ожидания в очереди к кассам

Цель: Проверить гипотезу о том, что среднее время ожидания в очередях не превышает 3:50 минуты, предложить меры по оптимизации.

Библиотеки: Python, pandas, matplotlib, scipy.

Методы:

- Бутстрапирование для оценки доверительных интервалов.
- Центральная предельная теорема для анализа больших выборок.
- Односторонний t-тест для проверки гипотезы.

Результаты:

- Установлено превышение среднего времени ожидания 3:50 минуты.
- Разработаны рекомендации по улучшению процесса обслуживания.

Данные: Две выборки (25 и 500 наблюдений) с учетом времени суток и дней недели.

Реализация градиентного бустинга с нуля

Цель: Реализовать алгоритм градиентного бустинга с нуля и сравнить с библиотечным решением.

Инструменты: Python

Методы:

- Реализация алгоритма на основе деревьев решений.
- Сравнение с sklearn.ensemble.GradientBoostingRegressor.
- Оценка качества с помощью метрик RMSE и MAE.

Подтверждена эффективность алгоритма, визуализирована сходимость ошибок и важность признаков