

РОМАН ФЕДОСЕЕВ

 github.com/RomanFedoseev  tg/FedoseevRS  fedoseeffsn2017@yandex.ru

ОБРАЗОВАНИЕ

Российская академия народного хозяйства и государственной службы

Москва, Россия

Бакалавр, "Цифровая экономика", институт ЭМИТ

2022 - н.в.

- **Учебные курсы:** Анализ данных на Python, Машинное обучение, Эконометрика, SQL, Теория вероятностей, Математическая статистика, Линейная алгебра, Глубокое обучение и т.д.
- **Дополнительные курсы:** Deep Learning School (семестр 1), Open Data Science - Classic ML models, DataCamp: Data Manipulation with pandas, Joining Data with pandas, Intermediate Python и т.д.

Школа аналитиков данных МТС Тета

Москва, Россия

Курсы повышения квалификации

2024 - н.в.

- Прошел конкурсный отбор в ежегодную программу МТС Тета, углубив экспертизу в А/В-тестировании, ML-моделировании и продуктовой аналитике с акцентом на прогнозирование временных рядов и поиск точек роста продукта.

ОПЫТ РАБОТЫ

Альфа-банк

Москва, Россия

Процессный дата аналитик / Python, Proceset, ClickHouse, Excel

Июль 2024 - Апрель 2025

- **Исследование бизнес-процессов:** провел анализ бизнес-процессов с помощью Process Mining: оптимизировал выдачу кредитов наличными (сокращение длительности на 7%) и выпуск дебетовых карт через интернет-анкету (сокращение длительности на 5%), выявив и устранив ключевые узкие места, разработав roadmap перехода к оптимальным dream-процессам.
- **Исследование операций сотрудников:** проанализировал работу премиум- и ипотечных менеджеров: провел стат. тесты для выявления различий между группами, предложил автоматизацию рутинных задач, что увеличило рейтинг менеджеров и улучшило клиентский опыт за счет устранения ключевых pain points в процессе взаимодействия.
- **Автоматизация расчета потенциала автоматизации:** разработал Python-бота (NumPy, Pandas) для вычисления внутренней метрики по подразделениям банка, что позволило сократить трудозатраты на 20 человеко-часов за две недели.
- **Автоматизация обработки Excel-отчетов:** разработал Python-скрипты для обработки данных (Pandas, numpy) для последующей загрузки в ETL-систему Proceset.

НАВЫКИ

- **Языки программирования:** Python, R
- **Инструменты работы с данными:** PostgreSQL / ClickHouse, Pandas, NumPy, Scipy, Excel
- **Машинное и глубокое обучение:** Scikit-learn, CatBoost, XGBoost, LightGBM, PyTorch
- **Визуализация данных:** Matplotlib, Seaborn, Plotly, Bokeh, GeoPandas, graphviz, PowerPoint, Figma, Canva
- **Инструменты и среды разработки:** Superset, Proceset, PowerBI, PL/SQL Developer, Jupyter Notebook, Colab
- **Дополнительно:** А/В-тестирование, Confluence, Microsoft Visio, LaTeX, Agile/Scrum, Miro (Kanban)
- **Языки:** Русский (родной), Английский (B1)

- **Научное исследование: прогнозирование волатильности фондового рынка с помощью методов ML.** Разработал модель машинного обучения для предсказания недельной реализованной волатильности на основе мультифакторных данных. Особое внимание уделено обработке мультиисточников и отбору признаков. **(смотреть решение...)**
- **Репликация классов PyTorch и мультирегрессия на данных о бриллиантах.** Репликация ключевых компонентов PyTorch вручную с последующей реализацией нейросети для одновременного предсказания массы (карат) и стоимости алмазов. **(смотреть решение...)**
- **Kaggle: детекция мошенничества в транзакциях.** В рамках школы МТС разработана модель обнаружения фрода, результат — 5 место среди участников. **(смотреть решение...)**
- **Хакатон «Цифровой прорыв» (ЦФО): анализ обратной связи студентов.** В составе команды разработан телеграм-бот, система сбора обратной связи и ML-модель анализа текстов. Результаты визуализированы в интерактивном дашборде. **(смотреть решение...)**
- **Кейс от VK: поиск точек роста VK Видео.** Проведён анализ пользовательских данных и сформулирован ряд гипотез роста. Применены методы статистического тестирования и инструменты визуализации. Предложены инициативы по улучшению продукта. **(смотреть решение...)**
- **ML-классификация преддиабета.** Полный цикл проекта: от EDA и балансировки классов до обучения и сравнения моделей классификации для выявления преддиабетного состояния. **(смотреть решение...)**
- **Гендерный разрыв в зарплатах: эконометрический анализ по регионам РФ.** Построена регрессионная модель на основе данных ВШЭ, оценены масштабы и детерминанты гендерного разрыва в оплате труда. **(смотреть решение...)**
- **Kaggle: Playground Series S4E12 – прогноз страховых премий.** В команде реализована модель soft voting (LightGBM и XGBoost), результат — топ-27% среди почти 2400 участников. **(смотреть решение...)**
- **Хакатон «Цифровой прорыв» (ЮФО): прогноз оттока клиентов.** Построена модель предсказания оттока на основе очищенных и дополненных (через ML) данных. Разработано финальное решение с высокой точностью. **(смотреть решение...)**
- **Kaggle: отток пользователей.** Разработана модель, продемонстрировавшая высокую производительность ($AUC = 0.8475$) на тестовой выборке. **(смотреть решение...)**
- **Кейс-чемпионат FU Case Club x S+Consulting.** Предложена стратегия выхода консалтинговой компании на международный рынок. Проект выполнен в составе команды из 4 человек. **(смотреть решение...)**
- **Научная работа: Детерминанты инвестиционной активности в РФ.** Определены ключевые детерминанты инвестиционного поведения, проведён обзор литературы. **(смотреть решение...)**