

ГЛУШКОВ ЕГОР | DATA SCIENTIST / ML ENGINEER

+7-922-917-14-31 | ✉ egorglushkov2014@yandex.ru | 📱 t.me/plextheflex | 🐙 github.com/Exp98

ОПЫТ РАБОТЫ

Цифровой двойник (системы поддержки решений, моделирование; dtwin.ru) Санкт-Петербург
Аналитик данных Март 2024 – Октябрь 2025

- Прогнозирование финансовых потоков и оптимизация портфеля заказов для металлургического комбината
 - Разработал модель *LightGBM* для прогноза задержки оплат заказов ($MAE = 1.74$ дня — точность сопоставима с экспертной)
 - Реализовал генетический алгоритм оптимизации портфеля заказов ($SMAPE = 5-15\%$) для достижения целевых KPI
 - Создал интерактивный прототип (*Shiny*) для расчёта сценариев и визуализации результатов
- Мониторинг и анализ рыночных и закупочных цен на строительные материалы для застройщика
 - Реализовал модели для анализа факторов ценообразования (*CatBoost*, *LightGBM*) и семантическое сопоставление записей (*RuSBERT* + *FAISS*) для унификации данных из разных источников
 - Автоматизировал сбор и анализ рыночных цен на строительные материалы (*ETL* на *Prefect*, отчетность на *Quarto*)
- Формирование системы прогнозов и отчётности по показателям социально-экономического развития (СЭР) субъектов РФ
 - Построил факторную модель ($PCA + HHT-EMD + ARIMA$) для прогноза 80+ показателей СЭР по 85 регионам ($MAPE = 4-20\%$)
- Развитие системы технического обслуживания и ремонта оборудования для энергетической корпорации
 - Спроектировал хранилище данных (*PostgreSQL*, *EAV* + *SCD2*) и автоматизировал загрузку и обновление таблиц объёмом 500M+ строк (*R* / *Python*, индексы и процедуры *PL/pgSQL*)

Smooth Drug Development (клинические исследования; smoothdd.com) Санкт-Петербург
Биостатистик Июль 2022 – Февраль 2024

- Провёл статистический анализ данных более чем 20 клинических исследований (*R*, *SQL*)
- Автоматизировал подготовку воспроизводимых статистических отчётов с валидацией данных
- Применял методы регрессионного и дисперсионного анализа, анализа выживаемости и статистические тесты для оценки эффективности и безопасности препаратов

ОБРАЗОВАНИЕ

Санкт-Петербургский государственный университет
Магистр, «Математическое обеспечение», фак-т Математико-механический; ср. балл 4.63 2020 – 2022
Бакалавр, «Программирование и информационные технологии», фак-т ПМ-ПУ; ср. балл 4.67 2016 – 2020

ПРОЕКТЫ

Дипломный проект, СПбГУ | *Предсказание кода Голланда по данным психометрических тестов личности*

- Разработал модульный ML-конвейер (*R*, *Python*, *Shiny*) для предсказания шести факторов личности по неполным данным других тестов
- Реализовал ансамбль *multioutput*-регрессоров с оптимизацией весов методом роя частиц (*PSO*), достигнув $C_{ind} = 11.66$ (выше опубликованных аналогов), проведя сравнение различных постановок задач и моделей
- Защитил работу как магистерскую ВКР, выступил на научной конференции; лучшее решение внедрено в лаборатории СПб ФИЦ РАН

НАВЫКИ

Языки	Python, R, SQL
ML и обработка данных	NumPy, Pandas, Polars, scikit-learn, CatBoost, LightGBM, PyTorch
Визуализация и прогнозирование	Matplotlib, Plotly, Shiny, SHAP, statsmodels, NeuralForecast
Инструменты и инфраструктура	PostgreSQL, ClickHouse, Git, Docker, Airflow, FastAPI, LLM
Сферы	Classical ML, Time Series, NLP