

Опыт

DS competitions: сентябрь 2020 – по настоящее время

- **Sibur challenge 2020 (2/206)** – провел тщательную очистку и подготовку данных о химическом составе сырья на заводе и построил поверх нескольких линейных моделей для предсказания изменения состава при транспортировке.
- **INGV - Volcanic Eruption Prediction (63/620)** – из сырых данных с сейсмографов рассчитал таблицу новых признаков и на ее основе построил ансамбль линейных моделей с учетом кросс-валидации.
- **Riiid! Answer Correctness Prediction (1859/3395)** – построил модель на основе градиентного бустинга, предсказывающую успехи учеников на обучающей онлайн платформе с учетом временной структуры данных.

Дополнительное образование

stepic

- Введение в Data Science и машинное обучение (Bioinformatics Institute)
- Интерактивный тренажер по SQL (ДВФУ)
- Основы статистики. Часть 1 (Bioinformatics Institute)
- Основы статистики. Часть 2 (Bioinformatics Institute)

coursera

- Математика и Python для анализа данных (Яндекс и МФТИ)
- Обучение на размеченных данных (Яндекс и МФТИ)
- Поиск структуры в данных (Яндекс и МФТИ)
- Построение выводов по данным (Яндекс и МФТИ)
- Прикладные задачи анализа данных (Яндекс и МФТИ)

Навыки

Python: pandas, numpy, scipy

ML: scikit-learn, xgboost, lightgbm, optuna, catboost

Визуализация: tableau, matplotlib, seaborn

SQL: выборки, агрегации, joins

Английский: чтение документации и форумов/статей о data science

Образование

Кубанский государственный аграрный университет, г. Краснодар

Бакалавр промышленного и гражданского строительства

О себе

С 2019 изучаю программирование, начал с Python, затем переключился на библиотеки и инструменты, связанные с анализом данных и машинным обучением.