

Курсовой проект вебинара «Видеокурс от Megafon + курсовой проект»





Последовательность действий

- 1. Загрузка данных:
- 1.1. оптимизация объема занимаемой памяти;
- 1.2. преобразование признаков 'buy_time' из типа int к временному ряду.
- 2. Exploratory Data Analysis:
- 2.1. проверка данных на пропуски;
- 2.2. исследование распределений;
- 2.3. исследование корреляций;
- 2.4. поиск дубликатов.
- 3. Сохранение и чтение модифицированных датасетов.
- 4. Исследование повторяющихся id.





Последовательность действий

- 5. Объединение датасетов test и features.
- 6. Построение модели XGBoost.
- 7. Проведение кросс-валидации.
- 8. Сохранение полученной модели.
- 9. Формирование индивидуальных предложений для абонентов на основе полученной модели.
- 10. Сохранение итоговых результатов.



Принцип составления предложений для абонентов

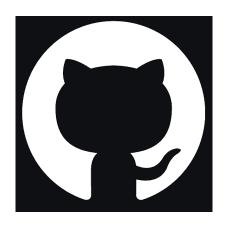
Для составления предложений для абонентов был использован следующий принцип:

- 1. Модель обучается на всей обучающей выборке.
- 2. Модели подаются на вход каждый класс подключенной услуги
- 3. По каждому варианту услуги считается вероятность подключения
- 4. Выбирается услуга, имеющая наибольшую вероятность подключения

Алгоритм можно улучшить, проведя Uplift-моделирование, что позволит избежать негативной реакции от части абонентов уже использующих «не базовые» тарифы или наборы услуг.



Ссылка на курсовой проект



https://github.com/lchimarouh/Megafon_course_project

