

Отчет по контесту

«Планирование продаж компаний ascott group»

Таскынов Ануар, 517 группа

Курс «Прикладные задачи анализа данных»

22 января 2018 г.

Даны: Информация о продажах. Все продажи поделены на группы по номеру филиала, номеру канала и id подгруппы товара (всего 108 комбинаций). Для таких комбинаций предоставлено количество продаж по неделям, дням. Также есть информация об активности на торговых точках. В обучающей выборке данные по продажам с января 2014 года. В тестовой содержатся 10 недель 2017 года.

Задача: Предсказать количество проданного товара из головного филиала ascott_group.

Функционал качества:

RMSE

Данных было недостаточно, так как в них не прослеживались закономерности.

Признаки:

- ① Количество троек (`idFilial`, `idSubGrp`, `KanalDB`). Несмотря на то, что в модели используется Catboost, данная фича немного улучшает качество.
- ② Категориальные фичи:
 - ① (`idFilial` + `KanalDB`),
 - ② (`idSubGrp` + `idFilial`).
- ③ `N wk`, `idFilial`, `KanalDB`, `idSubGrp`.

Модель:

100 Catboost'ов с разными сидами и обученные на разных выборках: 1 + 3 + 2.1 и 1 + 3 + 2.2.

Итог: **42211.26** на Public и **46066.53** на Private (30 место).