**Описание проекта и рекомендации.**

***В рамках разработки новых признаков были созданы следующие:***

1. deliveryDays - количество дней доставки после заказа
2. top\_sellerMark - отметка магазина с самым большим ассортиментом (предложением)
3. AddPrice - цена в сравнении с лучшей ценой витрины (AddPrice = FullPrice (полная цена с доставкой) / FullPrice (по минимальной цене в витрине).
4. MinDeliveryDaysMark – выбор предложения в витрине с минимальным сроком доставки

***Модель с предсказанием ранга (победы в витрине; качества прогнозирования рангов). Разделение датасета на обучающую и тестовую части, обучение модели:***

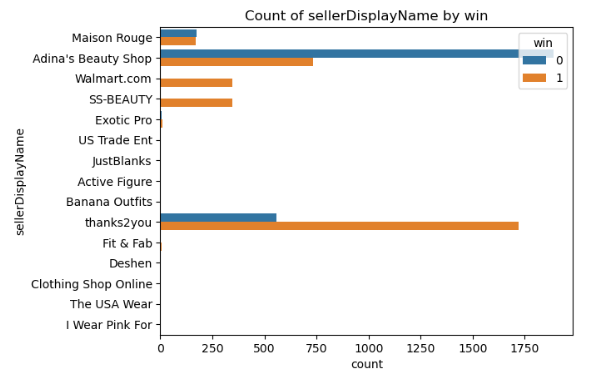
1. Модели были взяты lightgbm (LGBMRanker) и xgboost (XGBRanker). Предсказания модели можно интерпретировать как вероятности (чем меньше, тем лучше) того, что предложение будет релевантным (то есть будет на вершине витрины).
2. Деление датасета на обучающую и тестовую части было по группам (витринам) с помощью GroupShuffleSplit.
3. При обучении оценка качества усреднялась по 3-м выборкам.

***Мера оценки качества модели. Качество предложенной модели:***

1. Метрика качества ранжирования была взята ndcg.
2. Настройка моделей проводилась с помощью optuna.
3. Лучшая модель №2 получила по ndcg результат в 0,996 (по целевому признаку win) – данная модель максимально точно настроена на определение 0 ранга (победы). Также дополнительная оценка качества модели производилась по метрике, которая понятна будет для бизнеса «Доля выигранных витрин по предсказанию модели (bbwin=0)». Она получилась 0.9137, то есть в более 91% случаев модель правильно предсказывала ТОП витрины (предложение-победитель). А также по усредненному рангу по каждому предложению был создан прогнозный ТОП рангов продавцов:



Если посмотреть выигрыши на графике (win=0) в разрезе продавцов, то продавца Adina's Beauty Shop я бы также выделил на первое место в ранжировании, а Maison Rouge на второе.



1. Лучшая модель №1 получила результат по метрике ndcg в 0,990 (по целевому признаку order). Модель лучшая по качеству прогнозирования ранга на всех позициях в витрине. Но она уступает модели №2 по доли правильных предсказаний по выигрышам в витрине. ТОП продавцов по усредненному прогнозному рангу выглядит так:



***Код программы в файле ‘Best offer.ipynb’***

***Общие рекомендации:***

1. По минимальной цене в витрине продавалось почти 88% предложений (выигрыши).
2. Предложения, цена которых была выше на 24% от минимальной цены в витрине, почти не выигрывали (т.е. такая цена «непроходная»).
3. Важными для победы в витрине являются следующие факторы:

А) сравнение цены с минимальной ценой в витрине,

Б) сроки доставки,

В) большой ассортимент и, скорее всего, доверие к конкретному продавцу.

Важность признаков для прогнозирования по модели №2:

