**Проект: Обучение с учителем: качество модели**

**Описание проекта:** Интернет-магазин «В один клик» продаёт разные товары. Отчёт магазина за прошлый период показал, что активность покупателей начала снижаться. Привлекать новых клиентов уже не так эффективно, возможный выход — удерживать активность постоянных клиентов. Сделать это можно с помощью персонализированных предложений. Необходимо разработать решение, которое позволит персонализировать предложения постоянным клиентам, чтобы увеличить их покупательскую активность.

**Были обучены четыре модели:**

* KNeighborsClassifier(),
* DecisionTreeClassifier(),
* LogisticRegression(),
* SVC().

Для каждой из них подбраны гиперпараметры, устранён дисбаланс классов. Лучшая модель определена с помощью метрики roc\_auc.

**Лучшей моделью является LogisticRegression() с параметрами C=3, penalty='l1', random\_state=42, solver='liblinear'.**

**Оценены важности признаков для лучшей модели и построены графики важности с помощью метода SHAP.**

Менее всего значимы для модели признаки:

* Выручка за текущий месяц обсолютно не влияет на модель (вес признака 0,0);
* количество маркетинговых коммуникаций, категории товара, выручка за прошлые месяцы, время с момента регистрации, ошибки сервиса и рассылка покупателям дополнительных предложений о товаре влияют на работу модели незначительно (вес признаков менее 0,2);

Более всего значимы для модели признаки:

* Подавляющее влияние на модель оказывает количество просмотренных категорий товаров за визит (вес признака 0,69);
* Также немаловажную роль играют количество просмотренных страниц за визит и доля покупок товаров по акции (вес более 0,5).

**Выполнена сегментация покупателей,** с использование результатов моделирования и данных о прибыльности покупателей. Выбрана группа покупателей и предложены варианты увеличения её покупательской активности:

Покупательская активность не снижается, если подеррживаются на высоком уровне количество просмотренных категорий и страниц за визит. В то же время лояльность сохранят те, кто был активен в предыдущие месяцы. Немаловажен и тот факт, что **высокая доля купленных товаров по акции и большое количество неоплаченных товаров в корзине наоборот увеличивают вероятность снижения покупательской активности.**

Из всего вышеперечисленного можно сделать вывод, что **необходимо улучшать алгоритмы, отвечающие за предложение похожих товаров, а также товаров из смежных категорий.** Логично, что при большом выборе и правильном подборе нужных товаров повышается вероятность покупки. А вот большое количество акций ситуацию не спасают, а наоборот вредят покупательской активности. Возможно, клиенты относятся настороженно к акциям и скидкам, предполагая, что качество и полезность таких товаров низкое, если их не берут за свою цену.

Также справедливым будет такое предположение: если активно заинтересовывать клиентов новыми товарами и категориями, их активность будет помесячно высокой. Модель логично относит таких покупателей с высокой долей вероятности к тем, чья покупательская активность не снизится.