

Аналитический отчет об эффективности проведенной маркетинговой кампании и о факторах, способных повысить продажи.

Задачи:

1. определить эффективность маркетинговой компании;
2. выявить кластеры, на которые разбивается аудитория;
3. предложить методы работы с каждым кластером;
4. построить модель склонности клиента к покупке определённого товара.

Решение:

1. скидочная стратегия маркетинговой компании была неэффективна для увеличения объема покупателей, но эффективна для повышения среднего чека;
2. выявлено, что аудитория разбивается на 6 кластеров, но они представлены не равномерно;
3. предложены методы для эффективной работы с каждым кластером покупателей;
5. построена модель, позволяющая с высокой точностью предсказывать склонность клиента к покупке определённого товара.

Рекомендации:

1. провести дополнительное тестирование альтернативных стимулов, которые будут влиять на конверсию, пересмотреть ценовую стратегию для долгосрочного увеличения прибыли, рассмотреть инвестиции в дополнительные механизмы удержания клиентов;
2. увеличить выборку для более полной представленности каждого кластера и устранения дисбаланса между ними;
3. рассмотреть возможности таргетированной рекламы, регулярно обновлять ассортимент товаров, подчеркивая его качество и функциональность, внедрить персонализированную систему скидок и акций;
4. доработать модель, чтобы она учитывала дисбаланс кластеров, убрать из нее наименее значимые признаки.

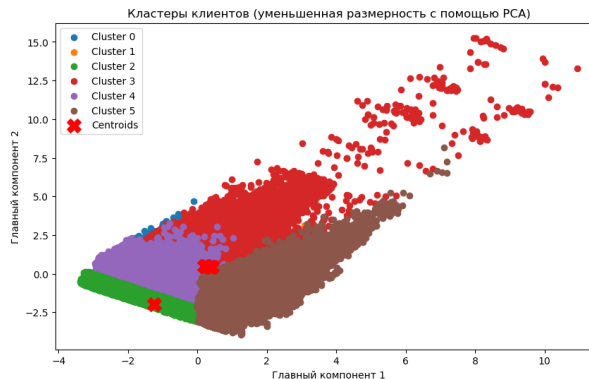
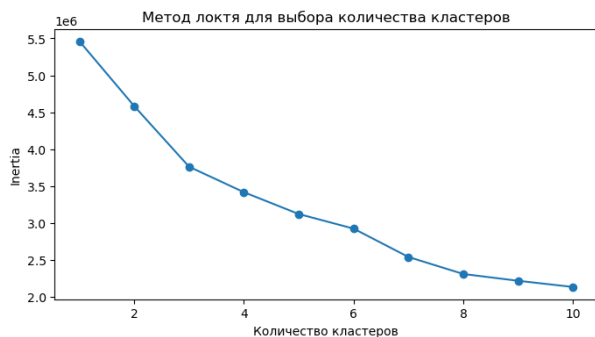
Исходные данные:

- личные данные клиентов (*personal_data*):
ID клиентов, их пол, возраст, образование, страна и город проживания;
- данные с персональными коэффициентами (*personal_data_coeffs*);
- данные о покупках (*purchases*):
ID клиентов, название товара, цвет, стоимость, гендерная принадлежность, скидка и дата покупки;
- данные об участии в маркетинговой компании (*ids_first_company_positive.txt*).

Данные нуждались в **улучшении качества**. Для решения этой задачи были

- восстановлены утерянные данные по части клиентов (пол клиентов);
- отфильтрованы данные только по стране с кодовым цифровым значением 32;
- обработаны пропуски в данных;
- выявлены и обработаны аномальные значения;
- данные о цвете товаров сделаны более однородными.

Был проведен кластерный анализ клиентов, который выявил 6 основных кластеров:



Кластер 0:

Средний возраст: 40.97 лет.

Средняя стоимость покупки: 4451.12.

Предпочтительные товары: Сумка FILA.

Влияние скидки: Большинство покупок совершается без скидки.

Рекомендации: Этот кластер состоит из покупателей среднего возраста, которые предпочитают аксессуары. Можно предложить новые модели сумок и акцентировать внимание на качестве и функциональности товаров.

Кластер 1:

Средний возраст: 41.68 лет.

Средняя стоимость покупки: 4734.31.

Предпочтительные товары: Брюки мужские Demix.

Влияние скидки: Практически равное распределение покупок со скидкой и без.

Рекомендации: Учитывая интерес к мужской одежде, стоит внедрить кампании, включающие скидки, и регулярно обновлять ассортимент спортивной одежды для мужчин.

Кластер 2:

Средний возраст: 42.76 лет.

Средняя стоимость покупки: 3150.17.

Предпочтительные товары: Футболка женская Demix.

Влияние скидки: Все покупки совершаются исключительно по скидке.

Рекомендации: Этот кластер сильно реагирует на скидки. Для их привлечения стоит регулярно предлагать акции на женскую одежду, особенно спортивные футболки.

Кластер 3:

Средний возраст: 41.21 года.

Средняя стоимость покупки: 7854.77.

Предпочтительные товары: Брюки мужские Demix.

Влияние скидки: Большинство покупок совершается без скидок.

Рекомендации: Это платежеспособная аудитория, предпочитающая мужскую одежду.

Им стоит предлагать премиальные коллекции и подчеркивать качество товаров.

Кластер 4:

Средний возраст: 43.68 года.

Средняя стоимость покупки: 4592.21.

Предпочтительные товары: Брюки женские Outventure.

Влияние скидки: Все покупки совершаются без скидок.

Рекомендации: Этот кластер ценит женскую одежду высокого качества. Можно предложить эксклюзивные коллекции или индивидуальные услуги, такие как персонализированные подборки.

Кластер 5:

Средний возраст: 23.37 года.

Средняя стоимость покупки: 5763.02.

Предпочтительные товары: Шорты мужские Demix.

Влияние скидки: Покупки равномерно распределяются между скидками и их отсутствием.

Рекомендации: Молодая аудитория, активно интересующаяся мужской одеждой. Для них можно использовать современные маркетинговые подходы, такие как таргетированная реклама в соцсетях, и акцент на новинках ассортимента.

Была построена **модель склонности клиента к покупке** определённого товара. Для ее построения были отфильтрованы данные только по жителям страны 32 города 188, в соответствии с пожеланиями заказчика.

Основные влияющие **факторы** (признаки для обучения модели):

- возраст клиента 'age',
- пол клиента 'gender',
- дата покупки 'dt',
- кластер клиентов 'cluster',
- гендерная принадлежность потенциальных покупателей товара 'product_sex'

Вывод по данным модели

Оценка точности модели:

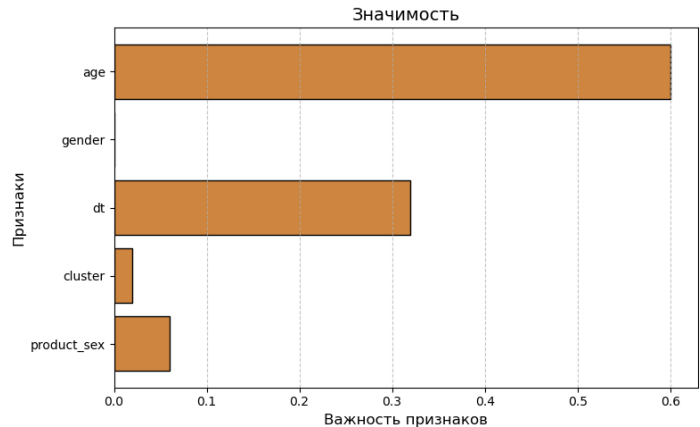
Модель достигает точности 1.00 на тестовых данных, что указывает на отличную предсказательную способность. Однако нулевая поддержка для класса 0 вызывает подозрение о наличии дисбаланса классов или недостаточной представленности класса в данных.

Кросс-валидация:

Стратифицированная кросс-валидация показывает почти идеальную точность (все значения выше 0.9999). Это подтверждает, что модель стабильна и надежна на разных подвыборках данных.

Значимость признаков:

Основной вклад вносит признак age (0.60). dt имеет вторую по значимости долю (0.32). Остальные признаки, такие как product_sex и cluster, играют менее важную роль, а gender вообще не влияет на предсказания.



Рекомендации:

Дисбаланс классов: Несмотря на высокую точность, стоит уделить внимание классу 0, так как он либо отсутствует в тестовых данных, либо модель его игнорирует. Возможно, потребуется увеличение выборки для класса 0 или использование методов, учитывающих дисбаланс классов (например, oversampling или weighted loss).

Оптимизация признаков: Можно рассмотреть исключение менее значимых признаков, таких как gender, для упрощения модели.

Заключение:

рекомендовано повторение маркетинговой компании, но необходимо большее таргетирование рекламы, а также предоставление персонализированных скидок с учетом особенностей поведения клиентов в рамках кластеров. Выборку клиентов можно осуществлять с помощью новой модели, но нужен промежуточный анализ в ходе компании, для проверки удовлетворенности клиентов и возможное влияние на лояльность, количество клиентов и размер среднего чека. Будет также полезным рассмотреть инвестиции в дополнительные механизмы удержания клиентов.