

УДК 004.85: 339.372

Рахматтилаев Бахтияр Кайратбекович

магистрант

im.baktyiar@kstu.kg

Картанова Асель Джумановна

к.ф.-м.н., доцент

asel.kartanova@kstu.kg

Сыдыкова Замира Абдыкеримовна

старший преподаватель

z.sadykova@kstu.kg

кафедра «Информационные системы в экономике»

Высшая школа экономики и бизнеса

Кыргызский государственный технический университет им. И. Раззакова

ИНСТРУМЕНТАРИЙ ИНТЕРНЕТ-МАРКЕТИНГА В ЗАДАЧАХ ПЕРСОНАЛИЗАЦИИ: АНАЛИТИКА ТРАНЗАКЦИОННЫХ ДАННЫХ И ПРОГНОСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

В данной статье рассматривается актуальная проблема повышения эффективности интернет-маркетинга малых и средних предприятий в эпоху цифровой трансформации экономики Кыргызстана. Выбор маркетинговых стратегий и поддержка управленческих решений в электронной коммерции напрямую зависит от инструментария интернет-маркетинга. Цифровые инструменты интернет-маркетинга позволяют проводить персонализацию во взаимодействии с клиентами путем сбора транзакционных данных. Бизнес-ориентированный анализ поведения клиентов интернет-магазина позволяет выявлять предпочтения клиентов, измеряя их покупательскую активность, и формировать будущие индивидуальные предложения. В работе предложена методология проведения бизнес-ориентированного анализа и получены результаты, позволяющие формулировать рекомендации и строить будущие маркетинговые стратегии для ведения эффективного бизнеса и получения максимальной прибыли объекту электронной коммерции.

Ключевые слова: электронная коммерция, эффективность интернет-маркетинга, МСБ, инструменты, персонализация, машинное обучение, аналитика данных, прогнозирование, модель, набор данных

Введение

В условиях цифровой трансформации экономики Кыргызской Республики (КР) выделены основные направления (рис. 1) для малого и среднего бизнеса (МСБ), такие как создание надежной цифровой среды для бизнеса, автоматизация бизнес-процессов и управление данными, развитие электронной коммерции и цифровых государственных сервисов, повышение финансовой доступности для малых и средних предприятий, поддержка культуры инноваций и цифрового предпринимательства.

Решение поставленных задач в совокупности создаст благоприятную почву для развития предприятий МСБ различных отраслей экономики, такие как торговля, сфера услуг, банковский сектор, электронная коммерция и другие. Ожидаемые результаты – это рост производительности и эффективности бизнеса, расширение рынка сбыта и клиентской базы, повышение прозрачности и качества управления, снижение издержек и операционных расходов и, самое главное, развитие инноваций и конкурентоспособности МСБ, и создание новых рабочих мест.

Рассмотрим электронную коммерцию, которая представляет собой торговлю продуктами или услугами через Интернет. Электронная коммерция опирается на такие технологии, как мобильная коммерция, электронный перевод средств, управление цепочками поставок, интернет-маркетинг, онлайн-обработка транзакций, электронный обмен данными, системы управления запасами и автоматизированные системы сбора данных.



Рисунок 1 – Основные направления цифровизации МСБ Кыргызстана

По данным кабинета министров КР прогнозируется, что сектор электронной коммерции Кыргызстана вырастет с 525 миллионов долларов США в 2025 году до 600 миллионов долларов США к 2028 году [1]. Потребители все больше оценивают удобство, разнообразие и доступность онлайн-платформ для совершения покупок. По данным исследований, более 80% людей в возрасте 20–30 лет предпочитают делать покупки через интернет-магазины и маркетплейсы. Ожидается, что этот тренд будет усиливаться по мере взросления этого поколения [2].

Одним из важных направлений развития предприятий МСБ является внедрение интернет-маркетинга для обеспечения доступа к широкой аудитории клиентов с относительно низкими затратами по сравнению с традиционным маркетингом. Для предприятий МСБ Кыргызстана, работающих в условиях ограниченного внутреннего рынка и высокой конкуренции, интернет-маркетинг становится эффективным инструментом расширения рынков сбыта и клиентской базы, повышения рыночной устойчивости. Интернет-маркетинг существенно расширит географические рамки деятельности, преодолевая территориальные ограничения традиционных каналов сбыта, и в особенности повысит конкурентоспособность малого и среднего бизнеса по сравнению с крупными компаниями.

В аналитическом обзоре TechTarget отмечено, что успешные примеры цифровой трансформации малого бизнеса показывают, что ключевыми факторами являются автоматизация процессов и внедрение управленческих и бизнес-решений, которые основываются на реальных данных и объективных метриках, а не на интуиции или предположениях [3].

Цифровые инструменты интернет-маркетинга позволят проводить персонализацию во взаимодействии с клиентами путем сбора транзакционных данных, а бизнес-ориентированный анализ поведения пользователей интернет-магазина, их предпочтений и покупательской активности позволит формировать будущие индивидуальные предложения.

Персонализация способствует повышению лояльности клиентов, росту процента потенциальных потребителей, что позволяет предприятию выстраивать долгосрочные и

плодотворные отношения с клиентами, повышая прибыль и оптимизируя маркетинговые стратегии. Таким образом, разработка стратегического инструмента интернет-маркетинга, необходимого для инновационного развития и эффективности деятельности предприятий МСБ, является актуальной проблемой.

Целью настоящего исследования является разработка и апробация методологии бизнес-ориентированного анализа данных интернет-магазина и моделей машинного обучения для формирования персонализированных маркетинговых стратегий и поддержки управленческих решений на предприятиях МСБ.

Объектом настоящего исследования является интернет-маркетинг предприятий МСБ в сфере электронной коммерции. Предмет – инструментарий бизнес-ориентированной аналитики данных, включающий методы сегментации и машинного обучения для прогнозирования отмены заказов и доходов, а также разработка рекомендаций на основе полученных результатов.

В работе [4] авторами данной статьи был проведен теоретический обзор состояния и динамики развития малого и среднего предпринимательства в КР. А также проанализирована ситуация с интернет-маркетингом и дана оценка благоприятных условий развития информационно-коммуникационных технологий, роста пользователей мобильной сети и интернета в целом, которые повышают привлекательность онлайн-торговли. Далее в рамках этой же статьи проведен предпроектный анализ эффективности внедрения интернет-маркетинга в деятельность предприятия МСБ, осуществляющего продажу потребительских товаров через интернет. Построена концептуальная модель информационной системы интернет-маркетинга предприятия в нотациях универсального языка моделирования UML. Одним из основных объектов этой системы является модуль «Аналитическая система». В данном исследовании раскрываются задачи, которые должна решать «Аналитическая система» для повышения эффективности интернет-маркетинга предприятия.

Аналитика больших данных (BDA – Big Data Analytics), как подчеркивается авторами в работе [5], дает МСБ инновационное преимущество, но требует выстраивание процессов интеграции знаний. Именно такой процесс – от сбора данных до маркетинговых решений, что особенно актуально для малых предприятий Кыргызстана рассмотрен в настоящей работе.

В работе авторы [6] подробно описали архитектуру веб-сайта электронной коммерции с системой онлайн-каталога. В нашем исследовании предлагается следующий шаг – это аналитический модуль, позволяющий не только управлять каталогом, но и персонализировать маркетинговые взаимодействия с клиентами. В статье [7] подчеркивается, что digital-маркетинг малых предприятий требует не только креативных стратегий, но и регулярного анализа поведенческих метрик. Данное исследование предоставляет конкретные метрики, такие как пожизненная ценность клиента (CLV-Customer Lifetime Value), RFM (Recency/давность, Frequency /частота, Monetary/сумма), вероятность отмены заказа для анализа поведения клиентов. Согласно принципам современного маркетинга [8], сегментация клиентской базы – основа для персонализации. В данной работе сегментация проведена не интуитивно, а с помощью RFM-метода и кластеризации, что соответствует современным подходам к принятию решений на основе данных.

Как рекомендуется в работе [9], любой ML-проект должен начинаться с четкого определения бизнес-цели. В нашем случае целью аналитики данных являются: снижение отмены заказов, повышение удержание клиентов и рост средней выручки. Соответственно для проведения машинного обучения выбраны три типа моделей: классификация, сегментация и регрессия. Основные манипуляции по аналитике данных и потенциальные выводы подробно описаны по разделам (табл. 1).

Таблица 1 – Основные манипуляции по аналитике данных

Анализ данных	Бизнес-аналитика	Машинное обучение
✓ Тенденции продаж ✓ во времени;	✓ Эффективность маркетинговых каналов;	✓ Прогнозирование отмены заказов;

✓ Поведение покупателей; ✓ Анализ эффективности продукта;	✓ Анализ влияния купонов; ✓ Тенденции в использовании способов оплаты	✓ Сегментация клиентов; ✓ Прогнозирование доходов
Потенциальные выводы:		
Какие продукты приносят наибольший доход? Какие источники реферальных ссылок привлекают больше всего клиентов? Какие факторы приводят к отмене заказов или возвратам? Как скидки влияют на покупательское поведение?		
Манипуляции с набором данных:		
Исследовательский анализ данных (EDA); Очистка и предварительная обработка данных; Разработка признаков; Визуализация (Matplotlib, Seaborn, Plotly); Прогностическое моделирование		

На рисунке 2 показан алгоритм действий в ML-системе, который представляет собой поэтапный бизнес-ориентированный анализ данных, содержащих информацию о поведении клиентов на сайте интернет-магазина, включающую данные о заказах товаров, о времени, проведенном на сайте и кликов мышкой, по прогнозированию покупок в электронной коммерции.

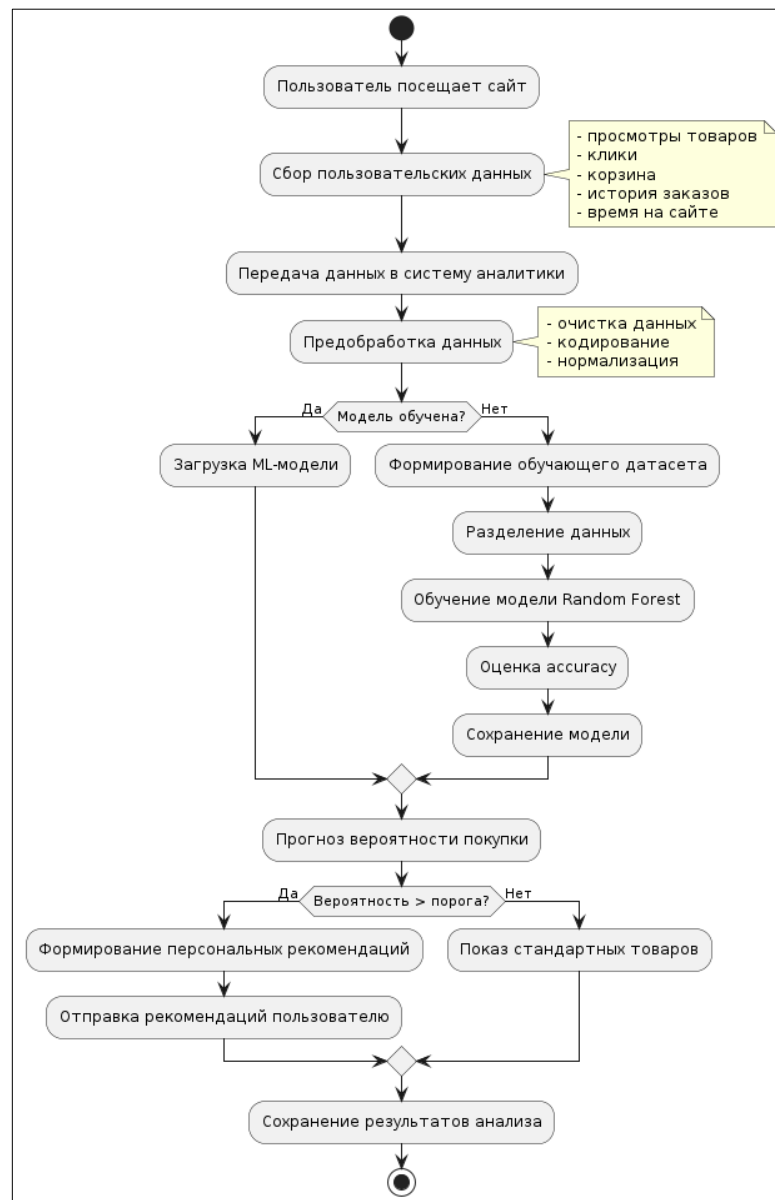


Рисунок 2 – ML-система прогнозирования в электронной коммерции

Материалы и методы

Рассмотрим набор данных, который представляет собой реалистичную симуляцию транзакций в электронной коммерции, отражающую поведение клиентов, модели заказов, предпочтения в способах оплаты и показатели продаж. Он предназначен для проектов анализа данных, машинного обучения и бизнес-аналитики, что делает его идеальным для студентов, аналитиков и специалистов. Используемый набор данных основан на публичном ноутбуке Kaggle [10], где продемонстрирован паплайн аналитики от начала до конца. Однако в отличие от исходного примера данная работа акцентирована на задаче персонализации интернет-маркетинга и адаптирует модели для выработки конкретных бизнес-решений.

Набор данных состоит из заказов в электронной коммерции (1200 заказов, 14 признаков, период с 2023 по 2025 гг.) и имитирует реальные платформы электронной коммерции, такие как Amazon, Shopify и Wildberries, основные характеристики набора данных (табл. 2).

Таблица 2 – Основные характеристики набора данных

№	Признак	Описание
1	OrderID	Уникальный идентификатор каждого заказа
2	Date	Дата заказа
3	CustomerID	Уникальный идентификатор клиента
4	Product	Приобретенный товар
5	Quantity	Количество приобретенных товаров
6	UnitPrice	Цена за единицу
7	ShippingAddress	Место доставки
8	PaymentMethod	Тип оплаты (кредитная карта, дебетовая карта, онлайн)
9	OrderStatus	Состояние заказа (доставлено, отправлено, отменено, возвращено)
10	TrackingNumber ID	Отслеживания отправления
11	ItemsInCart	Количество товаров, добавленных в корзину
12	CouponCode	Примененная скидка
13	ReferralSource	Канал привлечения клиентов
14	TotalPrice	Общая сумма транзакций

Подробнее исходные данные представляют собой историю заказов, содержащие сведения о заказанных товарах, покупателях, способах оплаты, о купонах на скидку, статусах заказов и датах оформления заказов, о месте доставки и отслеживания отправки товара, о каналах привлечения клиентов и об общей сумме транзакций.

В рамках настоящего исследования решены пять задач. Первая состояла в проведении исследовательского анализа транзакционных данных интернет-магазина с целью выявления временных закономерностей продаж, ключевых факторов формирования выручки и поведенческие характеристики клиентов. Вторая задача заключается в реализации сегментации клиентской базы для выделения однородных групп потребителей и формирование адресных маркетинговых рекомендаций по каждому сегменту. Третья задача предполагает построение и сравнительную оценку моделей бинарной классификации для прогнозирования отмен заказов на основе алгоритмов логистической регрессии, случайного леса и градиентного бустинга. Четвертая включает разработку регрессионных моделей прогнозирования выручки и оценку их точности по метрикам. Пятая задача состояла в формулировке практических рекомендаций по оптимизации интернет-маркетинга, непосредственно основанных на результатах аналитического конвейера, применительно к условиям электронной коммерции предприятий МСБ.

Алгоритм последовательного аналитического конвейера для интернет-магазина представлен на рисунке 3. Во-первых, была выполнена предобработка данных с целью приведения данных к виду, пригодному для анализа данных и машинного обучения.

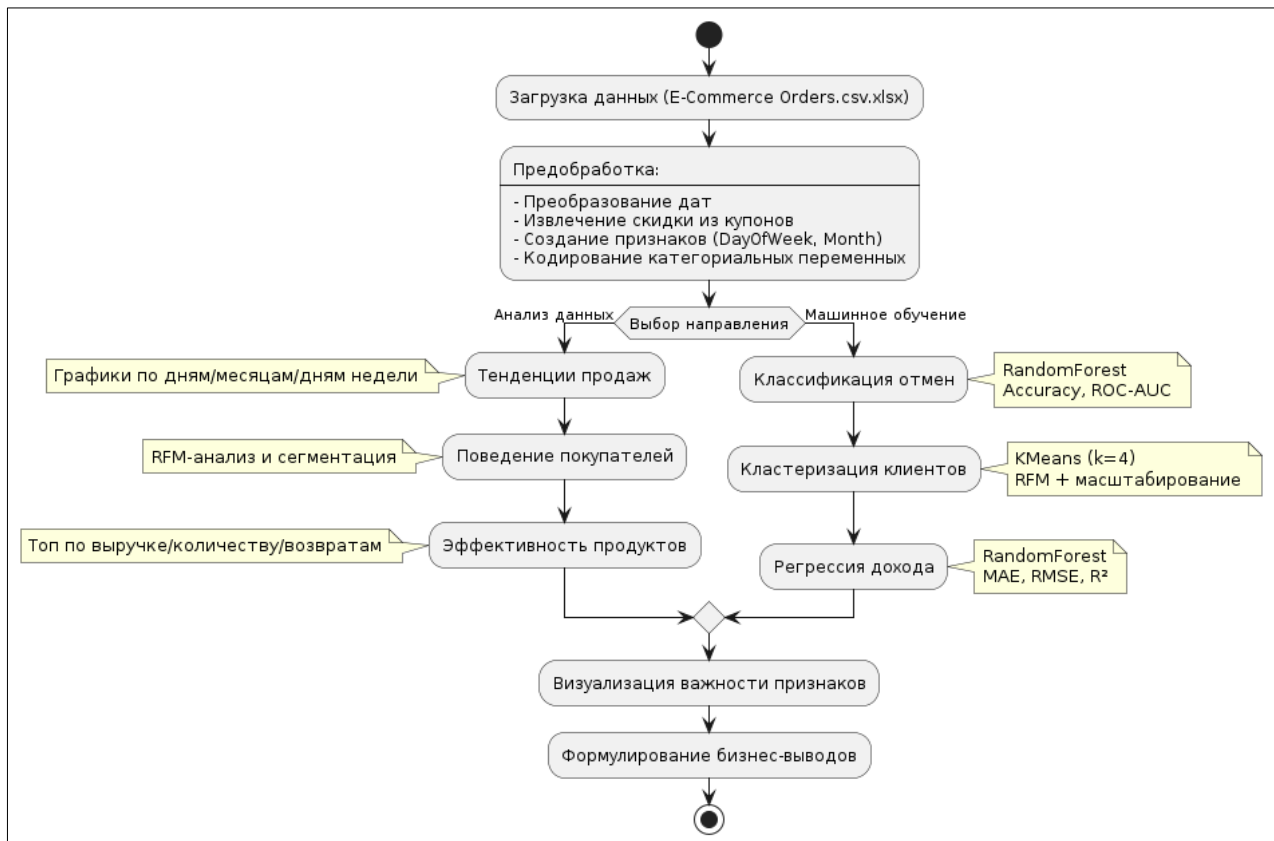


Рисунок 3 – Алгоритм конвейера аналитики данных

Предобработка данных включала удаление нерелевантных столбцов (TrackingNumber, ShippingAddress), преобразование поля Date в формат datetime с извлечением признаков Year, Month, DayOfWeek, извлечение числового значения скидки из поля CouponCode, формирование бинарных целевых переменных IsCancelled и IsReturned, а также кодирование категориальных признаков (Product, PaymentMethod, CouponCode, ReferralSource) методом LabelEncoder.

Данные разделены на обучающую и тестовую выборки в соотношении 80:20 (960 и 240 наблюдений соответственно), что соответствует стандартной практике в задачах контролируемого обучения. Для задачи классификации применялись три алгоритма: логистическая регрессия (baseline), случайный лес (Random Forest Classifier) и градиентный бустинг (XGBoost) [11].

Оценка качества производилась по метрикам: общая точность модели (accuracy), точность результатов (precision), полнота классов (recall), F1-мера (F1-score), площадь под кривой (ROC-AUC/Receiver Operating Characteristic - Area Under Curve), которая показывает качество ранжирования вероятностей и матрица ошибок. Применялся следующий стек технологий: язык программирования Python, библиотеки: Pandas, Scikit-learn, Plotly, Seaborn.

Результаты исследования и обсуждение

Проведя анализ тенденций продаж во времени для изучения временных закономерностей (рис. 4.), были построены графики ежедневной и помесечной выручки для исследования, в какие периоды продажи растут или падают, то есть изучить влияние сезонности на продажи.



Рисунок 4 – Панель мониторинга аналитики доходов

Ежедневные продажи продемонстрировали высокую волатильность, при этом явные пики приходились на отдельные даты, вероятно связанные с маркетинговыми акциями. Агрегированная по месяцам выручка показала заметный рост в конце календарного года, что указывает на наличие сезонного фактора. Анализ усредненной выручки по дням выявил, что наибольшие показатели приходятся на пятницу и субботу, что может служить основанием для концентрации рекламного бюджета в эти дни.

На следующем этапе исследовалось поведение покупателей (рис. 5, 6), которое было проанализировано с помощью RFM-методологии [12]. Для каждого клиента были рассчитаны три показателя: давность последнего заказа в днях, частота заказов и общая сумма заказа. По каждому из показателей клиентам присваивалась квартильная оценка от 1 до 4, после чего формировался составной RFM-балл. На основе этого балла были выделены четыре сегмента: чемпионы, в спячке, группа риска и новые клиенты.



Рисунок 5 – Панель мониторинга поведения клиентов



Рисунок 6 – Результаты сегментации клиентов

В результате получены портреты клиентов, показатели удержания клиентов для формирования последующих рекомендаций для персонализации.

На четвертом этапе провели анализ эффективности товара (рис. 7) с целью определения товара, которые приносят наибольший доход, маржинальность и имеют популярность. Провели расчет общего дохода по каждому товару, отсортированному по убыванию, и вычислили его кумулятивную долю. Получили категории (верхние 80% дохода), в (следующие 15%), с (последние 5%) и посчитали количество продаж и средний чек по каждому продукту. Оценили частоту возвратов/отмен по продуктам, и получен список ключевых товаров, на которые стоит делать акцент в маркетинге и закупках.

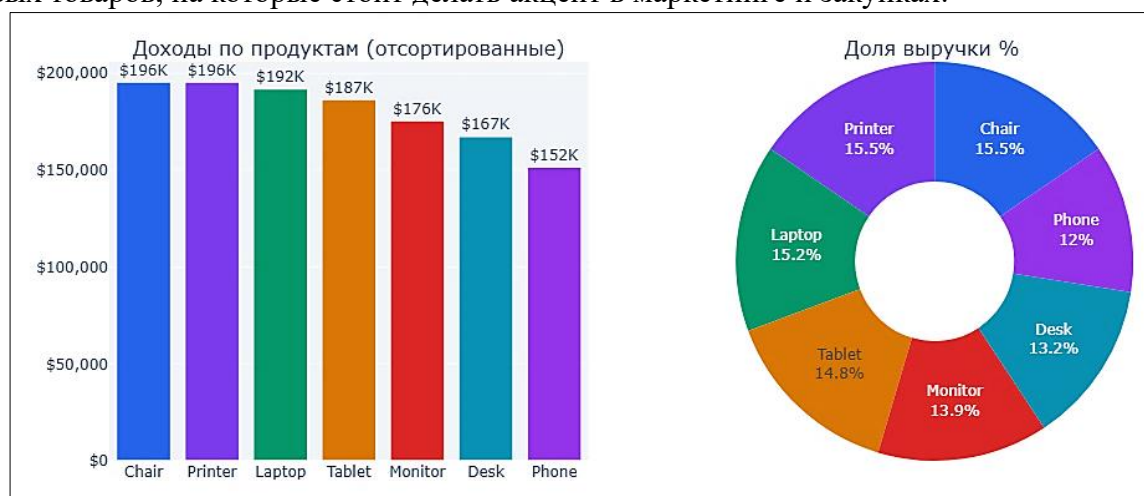


Рисунок 7 – Анализ доходов по продуктам/товарам

В рамках исследования разработаны три целевые модели, различающиеся по характеру зависимой переменной: бинарная классификация для идентификации фактов отмены взаимодействий (клиентов, транзакций), алгоритмы сегментация клиентской базы на основе многомерных признаков профилей, а также регрессионная оценка ожидаемого доходов.

Следующим этапом провели машинное обучение на наборе данных с целью прогнозирования отмены заказов, применив бинарную классификацию.

Получены результаты прогноза отмены заказа по трем моделям (рис. 8, 9). Результаты исследования показывают, что градиентный бустинг продемонстрировал наилучшие показатели AUC=0.534 среди рассмотренных трех моделей.

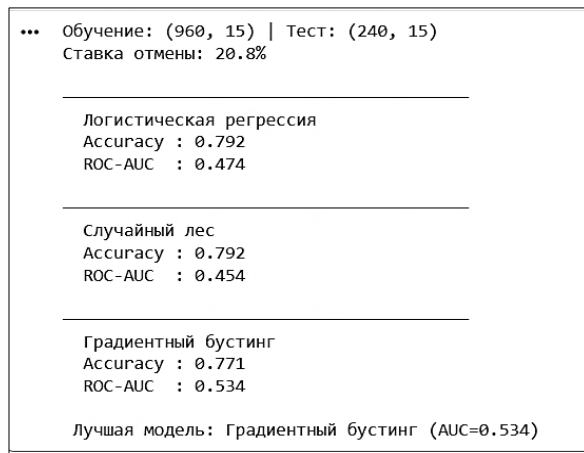


Рисунок 8 – Результаты прогноза отмены заказов по трем моделям

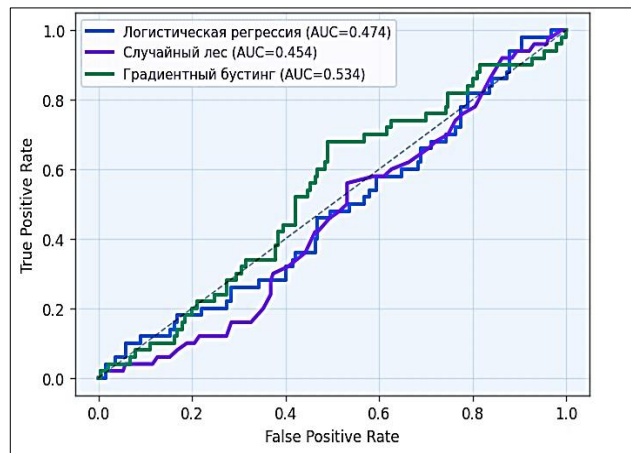


Рисунок 9 – Результаты прогноза отмены заказов, кривые ROC – три модели

Результаты моделирования прогнозирования отмены заказов показывают, что лучшая модель градиентного бустинга демонстрирует высокое значение точности ассурасу = 0.771, но показывает неэффективность модели для практического применения (полнота recall=0,04 и мера F1-score=0,07 для класса Cancelled). Высокое значение ассурасу объясняется дисбалансом классов (190 неотменённых против 50 отменённых заказов в тестовой выборке) и свидетельствует о практической непригодности модели для обнаружения риска отмены, (рис. 10, табл. 3).

Диаграмма важности признаков (рис.11) показывает факторы, которые наиболее сильно влияют на прогнозирование отмены заказов. Очевидно, наиболее важными являются UnitPrice (цена товара), AOV (средняя стоимость заказа), DaysSinceLast (количество дней с последнего заказа), средне важные – Month (месяц заказа), Product_enc(код товара), DayOfWeek (день недели), Payment_enc (код способа оплаты), слабо влияющие факторы – это Quantity (количество товара), HasCoupon (наличие купона) и IsWeekend (фактор выходного дня).

Корреляция между способом оплаты (наличные) и отменой заказа составила 0,34, а признак UnitPrice вошёл в топ-3 по важности в модели случайного леса, что согласуется с наиболее значимыми факторами для отмены заказа из работы [13], которыми являются цена товара и способ оплаты (наличные/при доставке) .

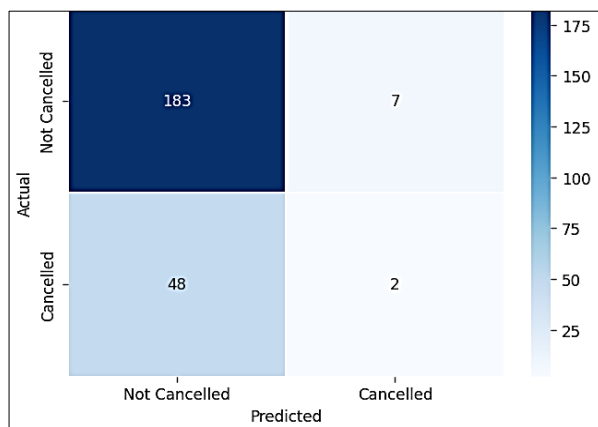


Рисунок 10 – Матрица ошибок, градиентный бустинг

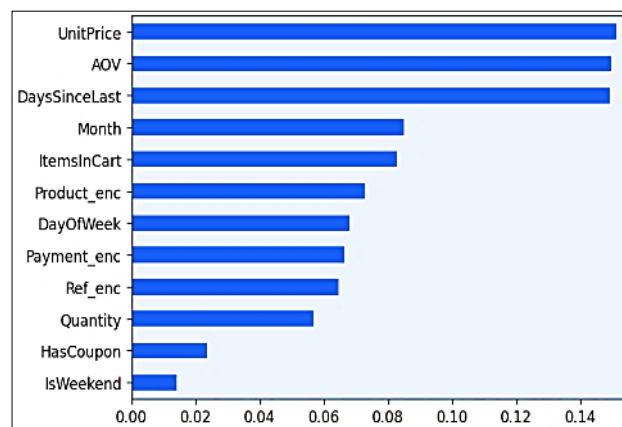


Рисунок 11 – Важность признаков, случайный лес

Таблица 3 – Отчет о классификации лучшей модели (градиентный бустинг)

	precision	recall	f1-score	support
Not Cancelled	0.79	0.96	0.87	190

Cancelled	0.22	0.04	0.07	50
accuracy	-	-	0.77	240
macro avg	0.51	0.50	0.47	240
weighted avg	0.67	0.77	0.70	240

Дисбаланс классов преодолеваем путем улучшения качества прогноза отмены заказа и производительности лучшей модели градиентного бустинга. Одним из основных способов борьбы с несбалансированными данными является увеличение выборки. Увеличение выборки заключается в добавлении дополнительных примеров класса-меньшинства, чтобы сделать его более представительным. Это можно сделать различными способами, например, путем дублирования существующих примеров, генерации синтетических данных или комбинирования этих методов.

В нашем случае дисбаланс классов можно решить путем применения весовых коэффициентов классов (`class_weight='balanced'`), методов SMOTE для оверсэмплинга (`oversampling`) миноритарного класса, а также снижения порога классификации с 0,5 до 0,3 для повышения чувствительности к отменам.

Для улучшения качества предсказания отмены заказа, особенно учитывая низкую полноту для класса Cancelled, мы применим технику SMOTE (Synthetic Minority Over-sampling Technique) – увеличение количества образцов миноритарного класса (отмененные заказы) и проведем настройку гиперпараметров для достижения наилучшей производительности этой модели, (рис. 12).

После применения SMOTE для балансировки классов: ROC-кривая (рис.13) для модели градиентного бустинга показывает значение $AUC=0.456$. Это ниже, чем у исходной модели ($AUC=0.534$), что указывает на снижение общей способности модели различать классы. Матрица ошибок (рис.14) показывает, что полнота (recall) для класса Cancelled значительно улучшилась (с 0.04 до 0.18), но при этом точность (precision) для этого класса снизилась (с 0.22 до 0.15) (табл. 4).

```

... Обучающие данные до SMOTE: (960, 15)
Обучающие данные после SMOTE: (1520, 15)
Распределение классов до SMOTE:
IsCancelled
0    760
1    200
Name: count, dtype: int64
Распределение классов после SMOTE:
IsCancelled
0    760
1    760
Name: count, dtype: int64

Производительность модели после SMOTE:
Accuracy : 0.617
ROC-AUC  : 0.456
    
```

Рисунок 12 – Результаты прогноза отмены заказов модель градиентного бустинга после SMOTE

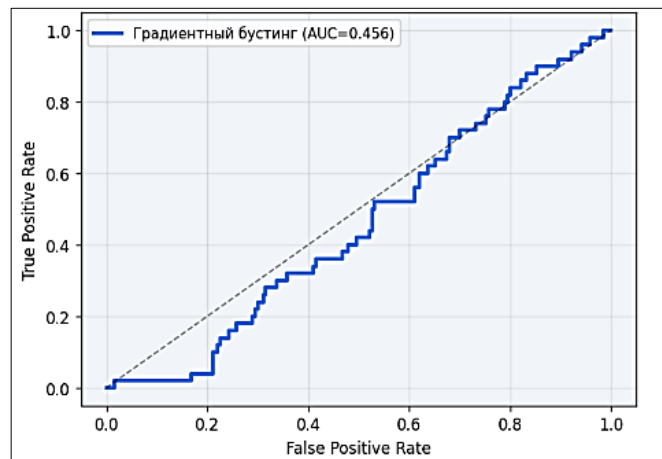


Рисунок 13 – Кривые ROC градиентный бустинг после SMOTE

Это означает, что модель теперь лучше обнаруживает отмененные заказы (рис. 14), но за счет увеличения числа ложных срабатываний. Важность признаков для модели градиентного бустинга после SMOTE (рис. 15) выделила AOV (средняя стоимость заказа), DaysSinceLast (количество дней с последнего заказа) и IsWeekend (является ли заказ выходным днем) как наиболее влиятельные.

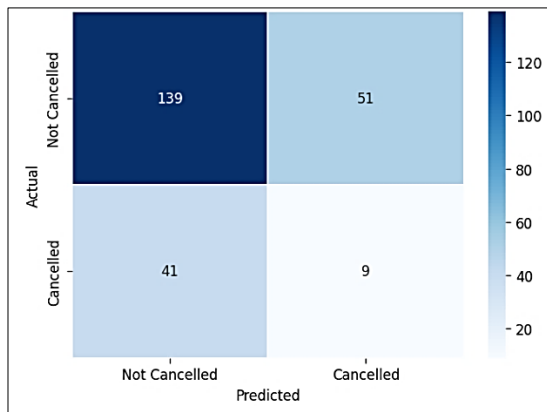


Рисунок 14 – Матрица ошибок, градиентный бустинг после SMOTE

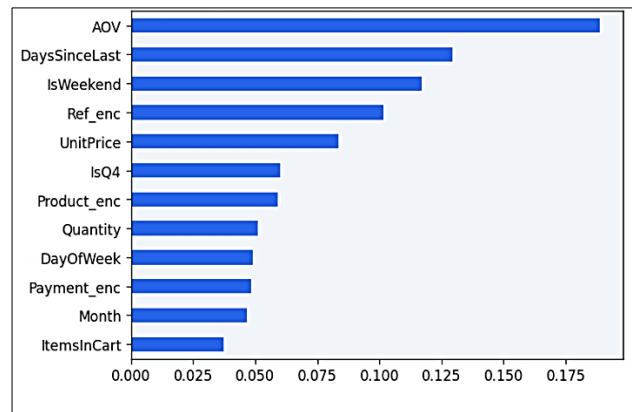


Рисунок 15 – Важность признаков, градиентный бустинг после SMOTE

Таблица 4 – Отчет о классификации модели градиентный бустинг после SMOTE

	precision	recall	f1-score	support
Not Cancelled	0.77	0.73	0.75	190
Cancelled	0.15	0.18	0.16	50
accuracy	-	-	0.62	240
macro avg	0.46	0.46	0.46	240
weighted avg	0.64	0.62	0.63	240

Таким образом, SMOTE помог модели стать более чувствительной к отмененным заказам, но общая производительность (Accuracy и ROC-AUC) снизилась. Это частый компромисс при работе с несбалансированными данными.

Следующим шагом рассмотрим решение задачи прогнозирования выручки во времени. В отличие от случайного разбиения временное разбиение позволяет оценить, насколько хорошо модель предсказывает будущее на основе прошлых данных.

Добавим сегменты клиентов и каналы продаж в качестве признаков. Для этого из даты извлекаем месяц, день недели и флаги выходных или 4-го квартала, а также категориальные данные (продукты, методы оплаты) переводим в числовой формат, добавляем сегменты клиентов из RFM-анализа. Далее проводим хронологическое разделение – данные сортируем по дате. Первые 80% данных используем для обучения, а остальные 20% данных – для тестирования. Это позволит спрогнозировать реальную ситуацию на основе прошлого. Обучение модели проводили с использованием мощного алгоритма GradientBoostingRegressor, который строит последовательность деревьев решений для минимизации ошибки прогноза. Для оценки адекватности модели рассчитали метрики (R^2 и средняя абсолютная ошибка MAE) и построили график, сравнивающий фактическую выручку с предсказанной на тестовом периоде (рис.16).

Главная особенность этого подхода в том, что модель пытается найти закономерности в сезонности и поведении сегментов клиентов для предсказания динамики продаж.

В результате проведенного анализа данных и прогнозирования можно сделать следующие выводы. Первое очевидно, что цена за единицу товара является самым сильным фактором, определяющим выручку (корреляция $> 0,85$). Стратегия премиальных продуктов окупается сполна. Четвертый квартал стабильно превосходит другие кварталы, подтверждая всплески спроса в праздничный сезон.

Постоянные клиенты в среднем генерируют в три раза больше пожизненной ценности клиента (CLV), чем новые клиенты – удержание более прибыльно, чем привлечение новых клиентов. Требования предоплаты за дорогостоящие заказы наличными/с оплатой при доставке гарантируют 30% отмен таких заказов. Добавление последовательности писем с подтверждением заказа в течение 1 часа предотвратят 10% общих отмен. Отметка покупателей, впервые приобретающих дорогостоящие товары, для проактивного обращения в службу поддержки предотвратят 15% отмен новых клиентов [13].

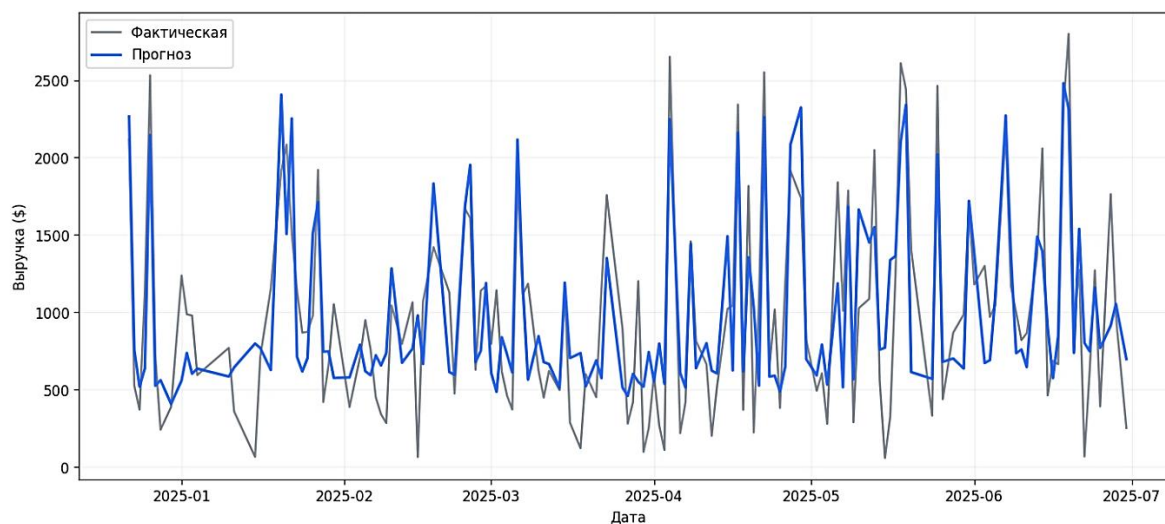


Рисунок 16 – Прогноз выручки во времени (Out-of-Time Test)

Новые клиенты отменяют заказы в два раза чаще, чем постоянные покупатели, поэтому необходимо улучшить процесс регистрации клиентов. Заказы без купона имеют несколько более высокий процент отмен, что указывает на ценовую чувствительность как на один из факторов. Самые ценные клиенты из сегмента «чемпионы», у которых высокая пожизненная ценность клиента (CLV), высокая частота обращений, низкий уровень оттока, таким клиентам необходимо VIP-обслуживание. На долю 20% лучших клиентов приходится около 65% от общего дохода (классический принцип Парето). Показатель «количество дней с момента последнего обращения» является сильным сигналом оттока, клиенты, неактивные более 90 дней, редко возвращаются.

По сути эффективности маркетинговых каналов, Google и электронная почта обеспечивают самый высокий доход на заказ, что указывает на высокую заинтересованность в покупке через эти каналы. Instagram обеспечивает объемы, но более низкий средний чек – он привлекает тех, кто ищет выгодные предложения, а не покупателей премиум-класса. Реферальные заказы имеют самый низкий уровень оттока – сарафанное радио привлекает клиентов более высокого качества.

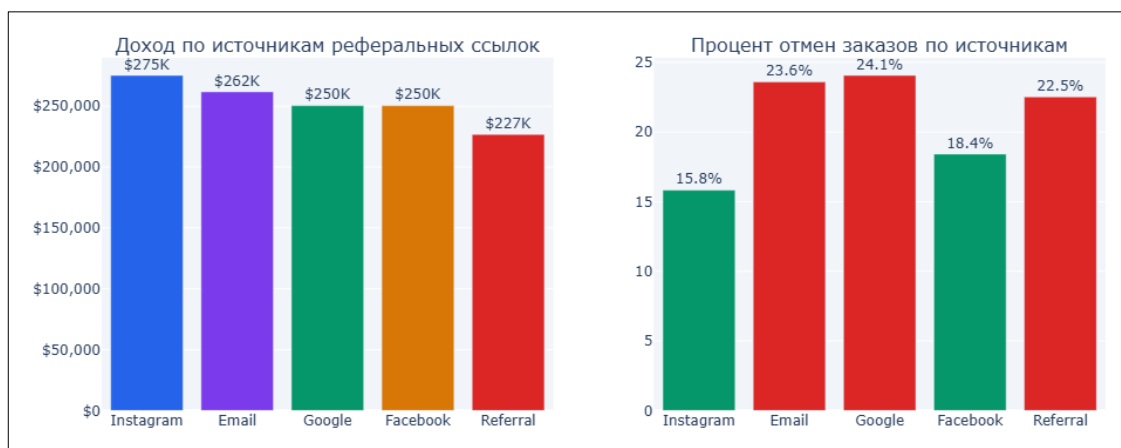


Рисунок 17 – Эффективность маркетинга по источникам реферальных ссылок

Практические рекомендации в ведении бизнеса – это комплектация товаров с высокой маржой (например, ноутбук + монитор), добавляет 8–12% от среднего чека. Запуск программы лояльности для постоянных клиентов обеспечит 5% повторных покупок. Расширение расходов на рекламу в Google и электронной почте (каналы с самым высоким ROI) обеспечат 10% квалифицированного трафика.

Заключение

Проведенный исследовательский анализ и машинное обучение моделей прогнозирования демонстрируют полный, готовый к использованию конвейер обработки данных на демонстрационном наборе данных электронной коммерции, который послужит инструментарием для разработки маркетинговых стратегий в электронной коммерции.

Эта поэтапная методология аналитики данных применима к аналогичным датасетам электронной коммерции предприятий МСБ КР и может быть адаптирована к реальным транзакционным данным при условии расширения набора признаков о типах клиентов и их поведенческих данных (демография клиентов, данные о просмотрах и геолокация клиентов).

В данной работе показана методология проведения бизнес-ориентированного анализа данных, раскрывающего информацию о доходах, клиентах, товарах, каналах продаж, с целью гибкого и продвинутого ведения эффективного интернет-маркетинга торгового предприятия, где критерием оптимальности является получение максимума прибыли.

Литература

1. [Электронный ресурс]. – Режим доступа: https://economist.kg/ekonomika/2025/10/30/obiem-rynka-eliektronnoi-kommiertsii-kyrgyzstana-dostigh-525-mdollarov/?utm_source=chatgpt.com (дата обращения: 01.06.2026)
2. Матненко, Н.Н. Развитие рынка электронной коммерции в условиях цифровизации экономики [Текст] / Н.Н. Матненко, Т.С. Сальникова, Ю.А. Мезяков // Экономические системы. – 2024. – Т. 17. – № 4. – С. 39–50. DOI 10.29030/2309-2076-2024-17-4-39-50.
3. 7 examples of digital transformation success in business [Электронный ресурс]. – Режим доступа: <https://www.techtarget.com/searchcio/tip/4-examples-of-digital-transformation-success-in-business> (дата обращения: 05.06.2026)
4. Рахматиллаев Б.К., Картанова А.Дж. Построение концептуальной модели бизнес-процессов интернет-маркетинга предприятия / Б.К. Рахматиллаев, А.Дж. Картанова // Известия Кыргызского государственного технического университета им. И. Раззакова. – 2026. – Т.2. – № 1 (77). – С. 359 – 371.
5. Cadden T., Weerawardena J., Cao G., Duan Y., McIvor R. (2023) Examining the role of big data and marketing analytics in SMEs innovation and competitive advantage. *Journal of Business Research*, 168, 114225.
6. Байгазиев М.С., Раззаков М.И., Каныбеков К.К. Разработка веб-сайта электронной коммерции с системой онлайн-каталога / М.С. Байгазиев, М.И. Раззаков, К.К. Каныбеков // Проблемы автоматизации и управления. – 2024. – № 1(49). – С. 41–51.
7. Катаев, А. В. Digital-маркетинг : учебное пособие / А. В. Катаев, Т. М. Катаева, И. А. Названова. – Ростов-на-Дону – Таганрог: Изд-во Южного федерального университета, 2020. – 161 с. – ISBN 978-5-9275-3437-1.
8. Котлер, Ф Маркетинг от А до Я. Основы маркетинга / Ф. Котлер // Альпина Паблишер. – 2024. – 707 с.
9. Амейзен Э. Создание приложений машинного обучения: от идеи к продукту. — СПб.: Питер, 2023. – 256 с.: ил. - (Серия «Библиотека программиста»). ISBN 978-5-4461-1773-4
10. E-Commerce Analytics: End-to-End Data Science [Электронный ресурс]. – Режим доступа: <https://www.kaggle.com/code/hammadansari7/e-commerce-analytics-end-to-end-data-science/notebook#F0%9F%A4%96-5.-Machine-Learning-Models> (дата обращения: 01.06.2026)
11. Груздев А. В. Предварительная подготовка данных в Python: Т. 1. Инструменты и валидация. – М.: ДМК Пресс, 2023. – 816 с.: ил.

12. RFM-анализ — что это такое и для чего нужен? [Электронный ресурс]. – Режим доступа: https://www.sberbank.ru/ru/s_m_business/pro_business/rfm-analyz-cto-takoe-i-dlya-chego-nuzhen (дата обращения: 01.06.2026)
13. Irmak T. (2025) Analysis of Order Cancellation Rates and Feature Weighting with Machine Learning: An E-Commerce Case Study. In Lecture Notes in Networks and Systems, Vol. 1531, pp. 763–770.