

УДК 004.62:668.5

Пересыпкин Александр Анатольевич

старший преподаватель,
кафедра финансов и учета,
Мелитопольский государственный университет,
г. Мелитополь, Российская Федерация,
e-mail: peresyarkin_oleksandr@mail.ru

Сосновская Ирина Леонидовна

старший преподаватель,
кафедра финансов и учета,
Мелитопольский государственный университет,
г. Мелитополь, Российская Федерация,
e-mail: melirina2@yandex.ru

ОПТИМИЗАЦИЯ БИЗНЕС-ПРОЦЕССОВ С ПОМОЩЬЮ МЕТОДОВ DATA MINING

Аннотация. В условиях стремительного роста объемов корпоративных данных традиционные подходы к управлению операционной эффективностью перестают обеспечивать требуемую скорость и точность решений. Целью статьи является систематизация и обоснование ключевых направлений применения инструментов интеллектуального анализа данных (Data Mining) для оптимизации сквозных бизнес-процессов. Рассмотрены методы классификации, кластеризации, поиска ассоциативных правил, анализа аномалий и прогнозирования временных рядов в контексте конкретных управленческих задач. Показано, как комбинация описательных и предсказательных моделей позволяет выявлять неочевидные узкие места, снижать операционные издержки и повышать качество обслуживания. Приведены примеры практической реализации в розничной торговле, логистике и банковской сфере, демонстрирующие экономический эффект от внедрения предложенных решений.

Ключевые слова: Data Mining, оптимизация бизнес-процессов, интеллектуальный анализ данных, машинное обучение, управление процессами, прогнозирование, кластеризация, ассоциативные правила, анализ аномалий.

Alexander A. Peresyarkin

Senior Lecturer,
Department of Finance and Accounting,
Melitopol State University,
Melitopol, Russian Federation,
e-mail: peresyarkin_oleksandr@mail.ru

Irina L. Sosnovskaya

Senior Lecturer,

Department of Finance and Accounting,

Melitopol State University,

Melitopol, Russian Federation,

e-mail: melirina2@yandex.ru

OPTIMIZATION OF BUSINESS PROCESSES USING DATA MINING METHODS

Abstract. With the rapid growth of corporate data volumes, traditional approaches to operational efficiency management are no longer providing the required speed and accuracy of solutions. The purpose of the article is to systematize and substantiate the key areas of application of data mining tools for optimizing end-to-end business processes. Methods of classification, clustering, search for associative rules, anomaly analysis and time series forecasting in the context of specific management tasks are considered. It is shown how the combination of descriptive and predictive models makes it possible to identify non-obvious bottlenecks, reduce transaction costs and improve the quality of service. Examples of practical implementation in retail, logistics, and banking are provided, demonstrating the economic effect of implementing the proposed solutions.

Keywords: Data Mining, business process optimization, data mining, machine learning, process management, forecasting, clustering, association rules, and anomaly analysis.

Современные организации функционируют в среде, насыщенной информационными потоками, порождаемыми транзакционными системами, IoT-устройствами, клиентскими взаимодействиями и внутренними журналами событий. Накопленные массивы содержат скрытые закономерности, способные радикально изменить представление об эффективности операционной модели. Однако сама по себе регистрация данных не дает конкурентных преимуществ — необходимо извлечение из них действенных знаний. Именно здесь на первый план выходит интеллектуальный анализ данных, предоставляющий математический аппарат для обнаружения паттернов, аномалий и зависимостей в слабо структурированных и многомерных выборках [1; 2].

Под оптимизацией бизнес-процессов в широком смысле понимается целенаправленное изменение последовательности операций, распределения ресурсов и правил принятия решений с целью минимизации затрат времени, финансов и труда при одновременном повышении качества результата. Традиционные техники улучшения, такие как Lean, Six Sigma и имитационное моделирование, опираются преимущественно на экспертные гипотезы и агрегированные показатели. Data Mining расширяет этот инструментарий, позволяя выявлять неочевидные связи, строить прогнозные модели и сегментировать объекты управления на ос-

нове объективных эмпирических свидетельств. Рост доступности вычислительных мощностей и зрелость алгоритмов машинного обучения последних пяти лет сделали внедрение соответствующих методик экономически оправданным не только в крупных корпорациях, но и в среднем бизнесе [3; 4].

С методологической точки зрения арсенал Data Mining принято разделять на две большие группы: описательные и предсказательные методы. Первые ориентированы на интерпретацию существующих данных и поиск скрытых структур, вторые — на построение моделей, способных предвидеть будущие события либо значения целевых переменных. В задачах оптимизации процессов обе группы используются взаимодополняемо, образуя аналитический контур непрерывного совершенствования [5].

Таблица 1

Классификация методов Data Mining, применяемых для оптимизации бизнес-процессов

Группа методов	Метод / алгоритм	Решаемая задача процесса	Пример внедрения
Описательные	Кластеризация (k-means, DBSCAN)	Сегментация клиентов, заказов, поставщиков	Выделение групп проблемных маршрутов доставки
Описательные	Поиск ассоциативных правил	Анализ цепочек событий, ведущих к сбоям или возвратам	Правила «действие → ошибка» в обработке заявок
Предсказательные	Классификация	Прогнозирование оттока, риска срыва поставки	Скоринг лояльности клиентов банка
Предсказательные	Регрессия временных рядов	Прогноз спроса, загрузки персонала	Планирование штатного расписания колл-центра
Описательные/предсказательные	Анализ аномалий	Выявление мошенничества, сбоев оборудования, узких мест	Мониторинг вибраций промышленных станков

Источник: [5].

Кластеризация представляет собой базовый описательный инструмент, решающий широкий спектр прикладных задач: сегментацию клиентов для дифференцированного обслуживания, типологизацию транзакций при выявлении узких мест, группировку поставщиков по уровню риска. Например, алгоритм k-средних, примененный к логам выполнения заказов, способен выделить кластеры «проблемных» маршрутов, характеризующихся аномально долгим временем комплектации. Дальнейший анализ центроидов таких кластеров дает руководителю фактологическую базу для перераспределения персонала или изменения складской топологии [1; 6].

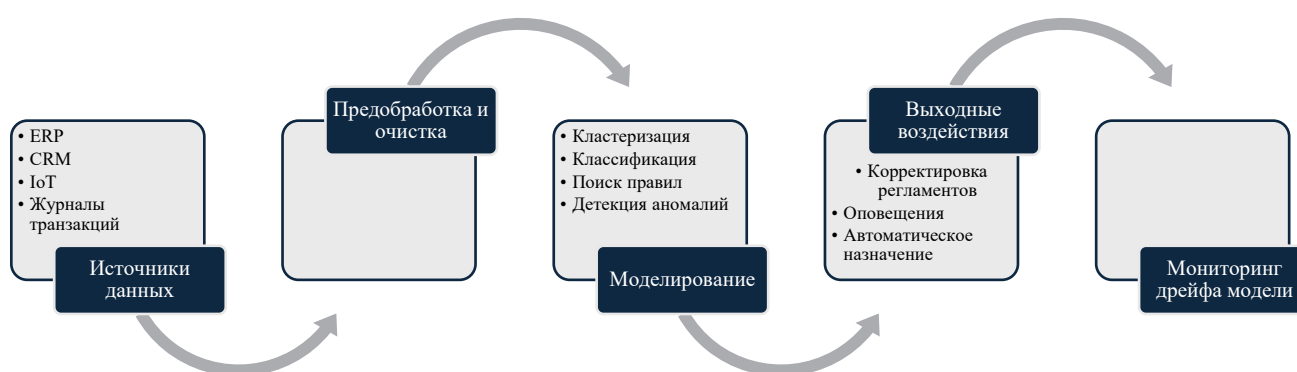
Поиск ассоциативных правил, исторически зарекомендовавший себя в анализе рыночной корзины, выходит далеко за пределы ритейла. В контексте бизнес-процессов данный метод позволяет обнаруживать устойчивые комбинации событий, предшествующих сбою, задержке или рекламации. Правило вида «если опе-

рация А выполнена с отклонением от регламента более чем на 20 %, то с вероятностью 78 % на следующем этапе возникает ошибка типа В» становится триггером для превентивной корректировки регламента. Алгоритмы Apriori и FP-Growth, реализованные в большинстве аналитических платформ, обеспечивают масштабируемый поиск таких правил в журналах ERP и CRM-систем [7].

Классификационные и регрессионные модели образуют ядро предсказательной аналитики. Логистическая регрессия, деревья решений, случайный лес и градиентный бустинг применяются для оценки вероятности наступления целевого события: оттока клиента, срыва срока поставки, отказа оборудования. Прогноз, генерируемый моделью в режиме реального времени, может быть встроен непосредственно в контур управления процессом — например, при автоматическом назначении приоритета заявке в службе технической поддержки на основе рассчитанного риска эскалации. Регрессионные модели временных рядов, включая SARIMA, активно используются для прогнозирования нагрузки на процессинговые центры, оптимизации штатного расписания и планирования закупок [2; 8].

Анализ аномалий заслуживает отдельного внимания, поскольку именно экстремальные отклонения в поведении системы часто сигнализируют о скрытых неэффективностях или мошеннических действиях. Методы на основе расстояния (к-ближайших соседей), плотности и изолирующего леса позволяют выявлять нетипичные транзакции в финансовых потоках, аномальные паттерны потребления ресурсов в производственных циклах и нестандартные маршруты согласования документов. Интеграция детектора аномалий в BPM-систему обеспечивает автоматическое информирование владельца процесса о подозрительных отклонениях до того, как они приведут к финансовым потерям [4; 9].

На рисунке представлена концептуальная схема аналитического конвейера, иллюстрирующая путь от регистрации сырых событий до управленческого воздействия.



Аналитический конвейер оптимизации бизнес-процессов
на основе Data Mining

Источник: [4].

Применение перечисленных методов к реальным бизнес-процессам целесообразно структурировать по функциональным областям, в каждой из которых Data Mining решает специфический класс задач.

В управлении цепями поставок доминирующей задачей выступает прогнозирование спроса. Избыток товарных запасов замораживает оборотный капитал, а дефицит ведет к упущенной выгоде и снижению лояльности. Ансамблевые модели, комбинирующие временные ряды с внешними факторами (макроэкономические индикаторы, погода, маркетинговая активность), показали в исследованиях 2022–2023 гг. снижение ошибки прогнозирования на 15–25 % по сравнению с экспоненциальным сглаживанием [5]. Параллельно кластеризация SKU по критериям стабильности продаж и маржинальности позволяет дифференцировать стратегию пополнения: для высокооборотчиваемых позиций устанавливается динамический страховой запас, а для редких — применяется модель «сборки под заказ». Результатом становится сокращение логистических издержек на 7–12 % без падения уровня сервиса [3].

Обслуживание клиентов и маркетинг — еще одна сфера, где Data Mining приносит измеримый эффект. Сегментация клиентской базы на основе поведенческих переменных дает основу для персонализированных коммуникаций. Однако более глубокий уровень оптимизации достигается при построении моделей оттока: идентификация клиентов с высокой вероятностью расторжения договора в ближайшие 30 дней позволяет инициировать проактивные удерживающие мероприятия. В банковской практике внедрение градиентного бустинга над сотнями признаков, включая транзакционную историю и обращения в службу поддержки, повысило точность прогноза оттока до AUC 0,89 и сократило коэффициент оттока на 3,2 процентных пункта в течение полугода [6; 9].

Анализ ассоциативных правил в розничной торговле не ограничивается размещением товаров на полке. Правила, извлеченные из данных о возвратах и жалобах, подсвечивают комбинации «продукт — канал продажи — регион», генерирующие наибольший объем обратного потока. В одном из реализованных проектов для сети бытовой техники было обнаружено, что продажи определенной категории товаров через мобильное приложение сопровождаются в 2,3 раза более высокой долей возвратов по причине «несоответствие ожиданиям» по сравнению с офлайн-каналом. Углубленный анализ выявил недостаточную детализацию описаний в мобильной версии, после исправления которой частота возвратов снизилась на 18 % [7].

Производственные процессы и техническое обслуживание все чаще опираются на предиктивную аналитику. Временные ряды показаний датчиков промышленного оборудования обрабатываются алгоритмами обнаружения аномалий, что позволяет перейти от планово-предупредительных ремонтов к обслуживанию по фактическому состоянию. Исследование, проведенное на базе предприятия тяжелого машиностроения, продемонстрировало сокращение внеплановых простоев на 27 % после внедрения модели, анализирующей вибрационные спектры в режиме реального времени [8]. В административных процессах аналогичный подход применяется к анализу длительности цикла: выявление транзакций,

время выполнения которых значительно отклоняется от медианы, помогает локализовать перегрузку конкретных исполнителей или несовершенство регламентов.

Таблица 2

Эффекты оптимизации бизнес-процессов с помощью Data Mining

Отрасль / процесс	Примененный метод	Ключевой показатель	Достигнутый эффект
Розничная торговля (возвраты)	Ассоциативные правила	Доля возвратов в мобильном канале	Снижение на 18 %
Банковский ритейл (отток)	Градиентный бустинг	Коэффициент оттока клиентов	Минус 3,2 п.п. за 6 мес.
Логистика (управление запасами)	Кластеризация SKU + регрессия	Логистические издержки	Сокращение на 7–12 %
Тяжелое машиностроение (ремонт)	Isolation Forest	Внеплановые простои оборудования	Снижение на 27 %
Телекоммуникации (удержание)	Ансамблевые модели	Точность прогноза оттока (AUC)	Рост до 0,89

Источник: [7; 8].

Для успешного встраивания Data Mining в практику управления процессами необходимо соблюдение ряда условий, выработанных за последние годы.

Модели машинного обучения чувствительны к пропускам, выбросам и несогласованности форматов. Рекомендуются развертывание отдельного слоя предобработки, включающего импутацию пропусков, нормализацию и кодирование категориальных признаков [2; 4]. Второе — интерпретируемость результатов. Бизнес-пользователь не обязан разбираться в устройстве градиентного бустинга, но должен понимать, какие именно факторы привели к тому или иному прогнозу. Методы SHAP и LIME, активно развиваемые с 2020 г., решают эту задачу, предоставляя декомпозицию вклада каждого признака в итоговое решение модели [5]. Третье — организационная готовность. Оптимизация на основе данных требует кросс-функционального взаимодействия: владельцы процессов формулируют гипотезы, аналитики строят модели, IT-служба обеспечивает поток данных, а руководство принимает решения на основе цифровых доказательств. Без зрелой культуры работы с данными даже самая точная модель останется лишь экспериментальным прототипом [3; 6].

Проведенный анализ подтверждает, что методы Data Mining перестали быть узкоспециализированным инструментом исследовательских лабораторий и превратились в практический актив операционного менеджмента. Сочетание кластеризации, поиска ассоциативных правил, классификации и обнаружения аномалий позволяет решать широкий спектр оптимизационных задач: от сокращения времени выполнения заказа до превентивного предотвращения инцидентов. Ключевым фактором успеха становится не выбор отдельного алгоритма, а построение сквозного аналитического конвейера, обеспечивающего постоянную циркуляцию данных от регистрации событий до внедрения управляющих воздей-

ствий. Дальнейшие исследования целесообразно направить на разработку адаптивных моделей, способных автоматически перестраиваться при изменении внешних условий, и на более тесную интеграцию Process Mining с предсказательными методами Data Mining для создания самооптимизирующихся бизнес-процессов.

Список использованной литературы

1. Гришин А.В. Интеллектуальный анализ данных в управлении операционной эффективностью предприятий / А.В. Гришин, Д.А. Новиков // Прикладная информатика. — 2021. — Т. 16, № 4. — С. 55–67.
2. Witten I.H. Data Mining: Practical Machine Learning Tools and Techniques / I.H. Witten, E. Frank, M.A. Hall. — 4th ed. — Cambridge: Morgan Kaufmann, 2020. — 664 p.
3. Сергеев Д.А. Управление цепями поставок на основе предиктивной аналитики: опыт розничных сетей / Д.А. Сергеев, В.П. Кузнецов // Логистика и управление цепями поставок. — 2022. — № 3. — С. 112–121.
4. Chen H. Business Intelligence and Analytics: From Big Data to Big Impact / H. Chen, R.H.L. Chiang, V.C. Storey // MIS Quarterly. — 2020. — Vol. 44, No. 4. — P. 1165–1188.
5. Provost F. Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking / F. Provost, T. Fawcett. — 2nd ed. — Sebastopol: O'Reilly Media, 2022. — 408 p.
6. Петрова А.В. Кластеризация клиентской базы и прогнозирование оттока в телекоммуникационной компании / А.В. Петрова // Маркетинг и маркетинговые исследования. — 2023. — № 2. — С. 88–99.
7. Ларионов В.Г. Применение ассоциативных правил для оптимизации процессов управления возвратами в электронной коммерции / В.Г. Ларионов, А.А. Семенов // Экономика и математические методы. — 2022. — Т. 58, № 1. — С. 102–114.
8. Гусев А.Н. Прогнозирование отказов промышленного оборудования методами машинного обучения / А.Н. Гусев // Научно-технический вестник информационных технологий, механики и оптики. — 2023. — Т. 23, № 5. — С. 712–720.
9. Семенов А.А. Обнаружение аномалий в финансовых транзакциях с использованием методов Data Mining / А.А. Семенов // Финансы и кредит. — 2024. — Т. 30, № 2. — С. 134–145.