

ЗАКУРДАЕВА Ж. Е.

**DATA MINING КАК ЧАСТЬ МЕТОДОЛОГИИ АНАЛИЗА
КЛИЕНТСКОЙ БАЗЫ ТЕЛЕКОММУНИКАЦИОННЫХ КОМПАНИЙ**

Аннотация. По мере развития цифровых технологий современные компании должны учитывать изменения в поведении потребителей. Статья рассматривает особенности применения технологии Data Mining для анализа клиентской базы на примере рынка телекоммуникационных услуг. Среди возможностей Data Mining – обеспечение информационной прозрачности и прогнозирования взаимоотношений с клиентами.

Ключевые слова: интеллектуальный анализ данных, Data Mining, массив данных, клиентская база данных, рынок телекоммуникационных услуг.

ZAKURDAEVA ZH. E

**DATA MINING AS PART OF CLIENT DATABASE ANALYSIS
OF TELECOMMUNICATIONS COMPANIES**

Abstract. As digital technologies advance, modern companies must take into account changes in consumer behavior. The article considers the features of using Data Mining technology to analyze the client base on the example of the telecommunications services market. Among the possibilities of Data Mining is the provision of information transparency and forecasting of customer relationships.

Key words: Data Mining, data array, client database, telecommunication services market.

Конкуренция в телекоммуникационной отрасли в последнее время становится более активной и сложной. Индустрия мобильных услуг быстро растет, и все больше людей пользуются новыми мобильными веб-услугами. Кроме того, огромные объемы данных все чаще выходят на передний план этой коммуникации в связи со взрывом данных, поступающих от широкого спектра технологий, и большим потреблением веб-контента [1].

Способами достижения конкурентного преимущества для компаний являются, например, увеличение доли рынка или инновации продуктов или услуг. Мобильные услуги включают не только голосовые и текстовые сообщения, но и предоставляют потребителям приложения, похожие на компьютерные, такие как просмотр документов, просмотр Интернета и мультимедиа, для использования и общения с другими пользователями. Социальные изменения происходят рекордно быстро, и индивидуальное общение с другими людьми стало более частым благодаря использованию мобильных социальных сетей. Эти модели поведения мобильных пользователей более сложны, чем модели поведения предыдущих компьютерно-опосредованных систем [2].

По мере дальнейшего развития индустрии мобильных услуг можно извлекать все больше пользы, понимая модели использования отдельных потребителей или конкретных групп. Изучение конкретных групп и их влияния из этой гигантской и сложной базы данных, является важной темой, которая привлекает все больше внимания. Традиционно люди интересуются мнением и советами своих сверстников относительно различных решений, которые они принимают [2].

Сбор данных происходит через различные информационные каналы, которые регистрируют поток информации о заинтересованности клиента в конкретном вопросе. Проблема для компаний заключается в том, чтобы знать, с каким клиентом, как и когда связаться [3]. Исходя из этого, изучение поведения и предпочтений потребителей становится важным направлением.

Применение методов интеллектуального анализа данных в области управления взаимоотношениями с клиентами является новой тенденцией. Интеллектуальный анализ данных (Data Mining) – это метод, который обнаруживает ранее неизвестные взаимосвязи в данных [7]. Data Mining анализирует большие наборы данных и ищет общие правила, предоставляя прогнозы и модели поведения на основе входных данных. К примеру, в маркетинге Data Mining помогает людям разделить своих клиентов на точный сегмент и проводить более эффективные маркетинговые кампании.

Также методы интеллектуального анализа данных позволяют помочь в процессах управления клиентом, начиная с привлечения потенциальных клиентов и заканчивая управлением отношениями с клиентом. Управление взаимоотношениями с клиентами – это бизнес-стратегия, направленная на оптимизацию доходов и прибыли при одновременном повышении удовлетворенности и лояльности клиентов [4]. Эта стратегия, ориентированная на клиента, предполагает сбор, анализ и изучение комплексной информации о взаимоотношениях, чтобы рационально распределить ресурсы в соответствии с потребностями, спросом и поведением клиента [5]. Таким образом, компании развивают свои навыки общения с клиентом. Кроме того, они могут сосредоточиться на коммерческой составляющей привлечения и удержания клиентов, а также на максимизации долгосрочной потребительской ценности [6].

С применением Data Mining первоначально проводится анализ собранных данных, чтобы понять, выполняются ли соответствующие требования. Таким образом можно выявить ошибки и несоответствия в данных. После обнаружения проблем создается механизм исправления и улучшения возникшего положения, при этом подходы к их решению должны быть гибкими и соответствовать реальности исследуемых данных и бизнеса. Необходимо исключить ошибочные или избыточные записи, преобразовать переменные, устраняя

несоответствия или облегчая их интерпретацию, среди прочих операций [3]. После определения значимых переменных необходимо выполнить прогнозирование, при котором применяются алгоритмы интеллектуального анализа данных, которые также способствуют организации обратной связи. Данная модель представляет собой автоматизированный механизм, который поможет изучить поведение клиентов, а также позволит им влиять на предложения рынка [3].

Data Mining может применяться в комплексе с каким-либо приложением, либо быть уже «встроенным» в него. Чтобы быть эффективным, приложение для добычи данных должно делать три вещи [1]. Во-первых, оно должно иметь доступ к общеорганизационным представлениям данных, а не к данным конкретных отделов. Часто данные организации дополняются данными из открытых источников или приобретенными данными. Полученная база данных называется хранилищем данных. В процессе интеграции данных приложение часто очищает данные - например, удаляет дубликаты, выводит недостающие значения (когда это возможно) и создает новые производные атрибуты. Во-вторых, приложение для добычи данных должно добывать информацию в хранилище. Наконец, оно должно организовать и представить добытую информацию таким образом, чтобы обеспечить принятие решений [5].

Системы, которые могут удовлетворить одно или несколько их требований, перечисленных выше, варьируются от коммерческих систем поддержки принятия решений, таких как Lotus 1-2-3 и Pilot Software's Lightship, до специализированных систем поддержки принятия решений и исполнительных информационных систем, таких как SAS/EIS. Например, Lotus 1-2-3 может получить доступ к данным из хранилища данных, позволить своим пользователям выполнять различные статистические операции и представлять результаты с помощью простой деловой графики. Общая цель каждой операции по принятию решений определяет тип информации, которую необходимо добыть, и способы организации добытой информации. Например, поставив цель выявить хороших потенциальных клиентов для взаимных фондов, менеджер по банковским счетам, упомянутый ранее, неявно указывает, что она хочет сегментировать базу данных клиентов банка на группы родственных клиентов - таких как городские, женатые, с двумя доходами, в возрасте от 30 лет, с низким уровнем риска, с высоким уровнем благосостояния - и определить уязвимость каждой группы к различным типам рекламных кампаний.

После создания хранилища данных процесс добычи данных состоит из четырех основных этапов: выбор данных, преобразование данных, добыча данных и интерпретация результатов. Хранилище данных содержит множество данных, причем не все из них необходимы для достижения каждой цели интеллектуального анализа данных. Первым шагом в процессе добычи данных является выбор целевых данных. Например, маркетинговые базы

данных содержат данные, описывающие покупки клиентов, демографические характеристики и предпочтения в образе жизни. Выбранные типы данных могут быть организованы в нескольких таблицах: во время выбора данных пользователю может потребоваться выполнить соединение таблиц. Более того, даже после выбора нужных таблиц базы данных для выявления полезной информации не всегда необходимо изучать содержимое всей таблицы. При определенных условиях и для определенных типов операций по добыче данных (например, при создании классификационной или регрессионной модели) обычно менее затратной операцией является выборка соответствующей таблицы, которая могла быть создана путем соединения других таблиц, а затем добыча только выборки.

Впоследствии пользователь обрабатывает преобразованные данные с помощью одной или нескольких методик для извлечения нужного типа информации. Например, для разработки точной символьной модели классификации, которая предсказывает, продлят ли подписчики журнала свою подписку, менеджеру по тиражам может понадобиться сначала использовать кластеризацию для сегментации базы данных подписчиков, а затем применить индукцию правил для автоматического создания модели классификации для каждого нужного кластера.

Таким образом, благодаря Data Mining, сочетание системы сбора информации о клиентах, системы выявления закономерностей в данных и системы построения аналитической отчетности даст возможность компаниям узнать лучше клиентскую базу и построить более эффективные взаимоотношения со своими клиентами. При отсутствии должной аналитики вся собранная информация не будет использоваться в полной мере и не позволит обратить данные о клиентах в знания о них, способные принести дополнительные доходы. Возможности большей информационной прозрачности и прогнозирования взаимоотношений с клиентами приводит все большее количество компаний к выводу о необходимости автоматизации процесса получения знаний о своих клиентах.

СПИСОК ЛИТЕРАТУРЫ

1. Chih-Ping C. Forming digital self and parasocial relationships on YouTube // *Journal Consum.* – 2016. – Vol. 16, Iss. 1. – P. 232–254.
2. Chen C. P., Weng J.-Y., Yang C.-S. & Tseng F.-M. Employing a data mining approach for identification of mobile opinion leaders and their content usage patterns in large telecommunications datasets // *Technological Forecasting and Social Change.* – 2018. – Vol. 130. – P. 88–98.
3. Espadinha-Cruz P., Fernandes A. Grilo, A. Lead management optimization using data mining: A case in the telecommunications sector // *Computers & Industrial Engineering.* – 2021. – Vol. 154. – P. 107–122.

4. Guifang G. & Youshi, H. Research on the application of data mining to customer relationship management in the mobile communication industry / Proceedings of 3rd IEEE international conference on computer science and information technology // ICCSIT. – 2010. – Vol. 2. – P. 597–599.
5. Guo F. & Qin, H. Data mining techniques for customer relationship management // Journal of Physics: Conference Series. – 2017. – Vol. 910, Iss. 1. – P. 1–6.
6. D’Haen J., Van Den Poel D., Thorleuchter D. & Benoit D. F. Integrating expert knowledge and multilingual web crawling data in a lead qualification system // Decision Support Systems. – 2016. – Vol. 82. – P. 69–78.
7. Закурдаева Ж. Е. Сравнительная характеристика современного инструментария для моделирования бизнес-процессов компании [Электронный ресурс] // Огарев-online. – 2021. – № 8. – Режим доступа: <http://journal.mrsu.ru/arts/sravnitelnaya-karakteristika-sovremennogo-instrumentariya-dlya-modelirovaniya-biznes-processov-kompanii> (дата обращения 15.05.2022).