УДК 004.75:004.455:658.818

ПРОГНОЗИРОВАНИЕ ПОТРЕБНОСТЕЙ КЛИЕНТОВ НА ОСНОВЕ КЛАСТЕРНОГО АНАЛИЗА

[1]Гордеев В.В., генеральный директор
[2]Столяров А.Д., аспирант
[1]ООО «АЭРОЛАБС»
[2]ФГБОУ ВО «Национальный исследовательский ядерный университет»

Аннотация. С ростом количества доступных для компаний данных, которые могут быть использованы для принятия решений растут с одной стороны возможности для их использования, с другой сложности их обработки. Данные собираются из различных систем и при этом в разных сферах бизнеса зачастую пытаются решить схожие задачи, связанные с прогнозированием потребностей клиентов, однако, везде для этого будут использоваться разные интеграции и везде будут доступны разные данные. Из-за этого крайне важной становится задача разработки универсальных подходов, которые без доработки могли бы использоваться в разных отраслях для решения поставленные бизнес-задач. Именно разработке такого универсального подхода для прогнозирования потребностей клиентов посвящена данная работа.
Ключевые слова: большие данные, кластерный анализ, рекомендательные системы, персонализированный маркетинг.

В условиях жестких санкций задача цифровой трансформации экономики и повышения темпов экономического развития страны стоит как никогда остро, поэтому необходимы иные подходы к управлению с использованием инновационных цифровых технологий, которые предлагают новые возможности для повышения эффективности управления [1]. В условиях быстрых изменений экономической среды только способность к инновациям в ответ на вызовы и проблемы позволит компаниям обеспечить свою устойчивость, гибкость и развитие, чтобы изменения не угрожали им, а предоставляли возможности для роста [2]. От качества управленческих решений в организации зависит ее результативность и эффективность. Клиентоориентированность становится общепринятой стратегией выживания на конкурентном рынке, где компании конкурируют на основе удобства и эмоционального взаимодействия, а CRM-системы являются необходимым средством управления взаимоотношениями с клиентами [3].

В условиях массовой цифровизации бизнеса встает задача эффективного использования данных о клиентах, которыми располагают компании. Первым уровнем использования данных о клиенте является внедрение CRM систем, используемых в ручном режиме для отслеживания персоналом компании информации об истории взаимодействии с клиентом и существующей о нем информации.

Вторым же уровнем использования данных о клиенте является автоматизация взаимодействия, при котором сценарии взаимодействия реализуются CRM системой [4]. Плюсы встраивания предиктивной аналитики в рабочий цикл организации связаны с применением интеллектуальной аналитики для получения оптимального решений, сведением к минимуму уровня неопределенности, своевременным реагированием на изменения в показателях различных видов бизнеса и точным управлением рисками [5]. Базовым сценарием реализации такого взаимодействия является автоматизированная генерация персонализированных предложений [6]. При этом данный сценарий может быть реализован как в виде системы поддержки сотрудника компании (в виде подсказок, какие предложения стоит сделать клиенту в процессе диалога), так и в виде непосредственной доставки предложений клиенту.

С развитием технологий накопление и обработка больших массивов данных стала ключевым инструментом для выстраивания маркетинговых стратегий организаций. Клиент любой компании сейчас окружен массивом цифровых решений, каждое из которых может использоваться для получения ценных крупиц информации, позволяющих точнее построить портрет каждого клиента. На этом этапе известную сложность представляет интеграция данных из этих информационных систем в единую базу данных. Однако, не меньшей проблемой после этого является эффективная и релевантная обработка собранных данных, вследствие которой у компании должна появляться информация, применимая для принятия решений и достижения экономических задач.

Такая обработка информации всегда нетривиальна, так как зависит от:

- задач организации;
- объема располагаемых данных;
- целостности данных;
- типа данных (качественные, количественные или смешанные);
- квалификации лиц, принимающих решения и требований к выводу обработанной информации (степени автоматизации).

В конечном итоге любой инструмент анализа данных требует существенной адаптации под каждую конкретную компанию, что вызывает сложности с созданием подобных инструментов анализа данных и ставить перед классической проблемой: достаточно адаптивный инструмент становится очень сложным в освоении, тогда как легкие в освоении инструменты дают пользователю крайне небольшие возможности для адаптации.

Описываемый подход к прогнозированию потребностей пользователей на основе обработки больших данных позволяет компании очень гибко настраивать обработку данных, при этом сохраняя простоту настройки.

Задача, которая ставилась при разработке подхода – обеспечить возможность идентификации потенциальных потребностей данного пользователя в некоторый момент его жизненного цикла. При этом пользователь может быть как совершенно новый (без истории взаимодействия с

нашей компанией), так и пользователь, о котором у компании присутствует уже значительный массив исторических данных. Такая постановка задача необходима прежде всего для автоматизированного определения наиболее подходящего товара или услуги, который компания на основе прогноза потребностей могла бы предложить данному клиенту на текущем этапе его жизненного цикла.

Прогнозирование потребностей клиентов осуществлялось в контексте генерации предложений дополнительных сервисов для клиента в процессе взаимодействия с компанией. Для прогнозирования использовались исторические данные о транзакциях, совершенных ранее клиентами (какие действия они совершали с предъявленными им объявлениями – просмотры, клики по объявлениями, покупки). Каждому из совершенных пользователем действий с конкретным предложением (рекламой) присваивался определенный балл (простой показ предложения считается за 1 балл, клик по предложению – за 2 балла, а покупка – за 10 баллов). Суммируя эти баллы по каждому из показов предложения данному клиенту, высчитывался средний рейтинг предложения дополнительной услуги у конкретного клиента.

Для того, чтобы при таком подходе спрогнозировать потребность определенного пользователя, достаточно было в базе данных найти пользователя с наиболее похожим на текущего профилем тэгов и сформировать такое предложение, рейтинг которого у похожего пользователя наивысший. Общий принцип такого прогнозирования на основе кластеризации представлен на рисунке 1:

Шаг 1: Присвоение людям массива тэгов на основе логических правил

Шаг 2: Расчет дополнительных критериев в виде рейтинга объявлений

Шаг 3: Поиск людей, близких по критериям и с нужным поведенческим паттерном

Рисунок 1 – Схема работы алгоритма кластеризации

При разработке модели прогнозирования в качестве одной из основных предпосылок бралось ограничение, что данные могут быть на входе любого типа – как качественные, так и количественные, так и смешанные и с любым типом данных система прогнозирования должна работать одинаково исправно.

Данный подход был опробован и показал высокую эффективность при прогнозировании дополнительных услуг для пассажиров авиакомпании на разных этапах их жизненного цикла. Для этого использовались следующие записи о пассажирах:

- PNR (Passenger Name Record – именная запись пассажира). Это запись в CRS (Computer Reservation System – компьютерная система резервирования), которая содержит информацию о маршруте полета для пассажира или группы пассажиров.

- ETLR (Electronic Ticket Lift Report) – представляет все использованные полетные купоны, т.е. измененные в финальный статус в течение отчетного периода.

В рассматриваемом примере существует отраслевая специфика, которая заключается в том, что все данные по отраслевым стандартам хранятся в простом текстовом формате и для их использования в базе данных разрабатывался специализированный разработчик. Кроме того, расположение данных в текстовых файлах хоть и стандартизировано в соответствии с отраслевыми стандартами, однако, без знания этих стандартов разобраться в структуре и расположении данных крайне сложно, что затрудняет потенциальную интеграцию.

Рассматриваемые данные содержат как количественные записи (например, число пассажиров или возраст), так и качественные (например, пол, город, страна) – таким образом, они являются отличной основой для демонстрации разработанного подхода к прогнозированию потребностей клиентов.

В рамках разработки подхода к обработке данных сначала была предпринята попытка применения классических методов кластерного анализа для разбиения клиентов на группы – с целью прогнозирования потребностей текущего клиента на основе имеющихся паттернов поведения других членов группы. Однако, применение классического кластерного анализа (k-means, сети Кохонена и другие), не дали интерпретируемых результатов. Кроме того, все опробованные методы имели существенные недостатки в части работы либо с качественными, либо с количественными данными.

В результате нами был разработан собственный подход к кластеризации [7] через присвоение клиентам ярлыков («тэгов» или «меток») на основе логических правил – например, такие метки:

- #alone: летает один (ни одного полета с кем-то в брони);
- #generationX: рожден с 1964 по 1983;
- #m: мужчина
- и так далее

Количество данных меток не ограничено. Их применение приводит фактически к разбиению базы данных на множество пересекающихся кластеров. При этом использование меток крайне удобно в части настройки системы кластеризации – администратору системы достаточно прописать собственные правила присвоения тэгов для получения разбиений, релевантных для его бизнеса.

Таким образом, получилось, что каждому из клиентов (пассажиров) поставлен в соответствие некий набор тэгов в зависимости от характеристик

этого пассажира и его поведения – фактически, некий «отпечаток» данного пассажира в нашей системе.

Главным преимуществом предлагаемого подхода является его прозрачность для компании (в отличии от искусственного интеллекта, являющегося для пользователя «черным ящиком») и возможность гибко настраивать логику – не только в части присвоения меток, но и в части подходов к обработке данных. Например, для каждого пользователя считать не совокупный рейтинг предложения, а средний, либо менять весовые коэффициенты у воронки продаж и даже добавлять новые уровни в воронку продаж.

Инновационная активность персонала работы с данными о клиентах дает возможность для формирования активного инновационного развития, внедрения новых цифровых технологий и цифровых бизнес-моделей компаний [8]. Следует отметить, что изучение нового цифрового инструментария и управление взаимоотношениями с клиентами в динамическом режиме требует комплексных изменений в организации и повышения уровня знаний и компетенций ее сотрудников. Персонализированное обучение в организации должно проводиться в рамках системного подхода [9], внедряться на постоянной основе и стать принципиально новым инструментом для регулярного определения индивидуальных компетенций сотрудников и формирования программ обучения и развития персонала на основе обратной связи и анализа результатов.

Список литературы:

1. Абрамов В.И., Андреев В.Д. Проблемы и перспективы цифровой трансформации государственного и муниципального управления в регионе (на примере Кемеровской области) // Ars Administrandi. Искусство управления. – 2022. – Т. 14, № 4. – С. 667-700. – DOI 10.17072/2218-9173-2022-4-667-700. – EDN CTWMYG.

2. Абрамов В.И., Гордеев В.В., Столяров А.Д. Создание региональных бизнес-экосистем на основе цифровых профилей клиентов и омниканальных коммуникаций // Экономика, предпринимательство и право. – 2023. – Том 13. – № 5. – doi: 10.18334/epp.13.5.117670

3. Абрамов В.И., Абрамов И.В., Поливанов К.В., Семенков К.Ю. Цифровая трансформация системы управления отношениями с клиентами // Вопросы инновационной экономики. – 2023. – Том 13. – № 1. – doi: 10.18334/vinec.13.1.117051.

4. Абрамов В.И., Чуркин Д.А. Предиктивная аналитика взаимоотношений с клиентами как метод адаптации компании к изменениям и повышения ценности предложения // Экономика, предпринимательство и право. – 2022. – Т. 12. – № 6. – С. 1709–1722. – DOI 10.18334/epp.12.6.114842. – EDN GPPFPW.

5. Абрамов В.И., Чуркин Д.А. Оценка уровня зрелости системы управления взаимоотношениями с клиентами // Вестник университета. – 2022. – № 12. – С. 5–13. – DOI 10.26425/1816-4277-2022-12-5-13. – EDN FSACWD.

6. Столяров А.Д., Гордеев В.В., Абрамов В.И. Модель модуля динамической генерации персональных предложений дополнительных услуг для пассажиров авиакомпаний // Экономика и управление. – 2023. – Т. 29. – № 3. – С. 335–344. http://doi.org/10.35854/1998-1627-2023-3-335-344

7. Гордеев В. В., Столяров А.Д., Абрамов В.И. Технология кластеризации и генерации персонализированных торговых предложений для пассажирских авиаперевозок // Современные наукоемкие технологии. – 2023. – № 4. – С.34-41.

8. Шпак Ю.О., Абрамов В.И. Управление инновационной активностью персонала при цифровой трансформации компаний малого и среднего бизнеса // Вестник Кемеровского государственного университета. Серия: Политические, социологические и экономические науки. – 2022. – Т. 7, № 1(23). – С. 115-124. – DOI 10.21603/2500-3372-2022-7-1-115-124. – EDN DFKAFR.

9. Абрамов В.И., Глухова Е.В., Семенков К.Ю. Цифровая трансформация системы развития и обучения персонала предприятий // Лидерство и менеджмент. – 2023. – Том 10. – № 1. - С. 189-202 – doi: 10.18334/lim.10.1.117182.