

СТАТЬИ

УДК 004.9

DOI 10.17513/snt.39624

**МЕТОДИКА ПРОГНОЗИРОВАНИЯ ПОТРЕБНОСТЕЙ КЛИЕНТОВ
БИЗНЕС-ЭКОСИСТЕМ НА ОСНОВЕ КЛАСТЕРНОГО АНАЛИЗА****¹Абрамов В.И., ²Гордеев В.В., ¹Столяров А.Д.**¹ФГАОУ ВО «Национальный исследовательский ядерный университет «МИФИ», Москва,

e-mail: viabramov@mephi.ru, mr.alexst@gmail.com;

²ООО «АЭРОЛАБС», Москва, e-mail: v.gordeev@aerolabs.aero

В статье представлены результаты научно-исследовательской работы, направленной на разработку нового подхода к прогнозированию потребностей клиентов любых бизнес-экосистем с использованием кластерного анализа. Развивающийся информационный рынок предоставляет предприятиям возможность беспрецедентного доступа к данным о потребителях. Эта информация может быть ценной по многим причинам, в частности потому, что она может предоставить сведения о потребительских предпочтениях и готовности платить, поэтому растущая доступность потребительских данных порождает новые требования к их использованию и регулированию. Обсуждаются предпосылки и актуальность разработанного решения, типы данных, использовавшихся при создании программы, и общее содержание самого метода прогнозирования потребностей, включая применяемые приемы кластерного анализа, использование данных об активности клиентов – воронки продаж, и расчет сходства клиентов для проективных расчетов потребностей. На основе описываемых подходов была разработана методика прогнозирования потребностей клиентов, прозрачная для пользователя и позволяющая гибко выполнять настройку, работающая как с качественными, так и с количественными данными. Отличительной особенностью представленного здесь метода является его прозрачность, что важно для повышения доверия со стороны пользователей, а также то, что для выполнения расчетов, связанных с прогнозированием спроса, требуется небольшая инфраструктура.

Ключевые слова: большие данные, кластерный анализ, рекомендательные системы, персонализированный маркетинг, прогнозирование спроса

**METHODOLOGY FOR FORECASTING THE NEEDS OF BUSINESS
ECOSYSTEM CLIENTS BASED ON CLUSTER ANALYSIS****¹Abramov V.I., ²Gordeev V.V., ¹Stolyarov A.D.**¹National Research Nuclear University MEPHI, Moscow,

e-mail: viabramov@mephi.ru, mr.alexst@gmail.com;

²Aerolabs LLC, Moscow, e-mail: v.gordeev@aerolabs.aero

The article presents the results of research work aimed at developing a new approach to forecasting the needs of customers of any business ecosystem using cluster analysis. The emerging information marketplace provides businesses with unprecedented access to consumer data. This information can be valuable for many reasons, in particular because it can provide insights into consumer preferences and willingness to pay, so the growing availability of consumer data creates new requirements for their use and regulation. The background and relevance of the developed solution, the types of data used to create the program, and the general content of the needs forecasting method itself are discussed, including the cluster analysis techniques used, the use of customer activity data – sales funnels, and the calculation of customer similarity for projective needs calculations. On the basis of the described approaches, a customer needs forecasting technique was developed that is transparent to the user and allows flexible configuration, working with both qualitative and quantitative data. The distinguishing feature of the method presented here is its transparency, which is important for increasing user confidence, and the fact that little infrastructure is required to perform calculations related to demand forecasting.

Keywords: big data, cluster analysis, recommender systems, personalized marketing, demand planning

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. 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 предоставля-

ли возможности для роста [1]. 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 [2].

В наши дни, в эпоху становления шестого технологического уклада и во времена четвертой промышленной революции, ключевым фактором успеха является информа-

ция, и эта информация представляет собой, как правило, большие массивы данных, собранные в основном в неструктурированном виде. [3]. Согласно исследованию [4], компании должны адаптироваться к технологическим изменениям, иначе они не выживут в долгосрочной перспективе, поэтому предприятия должны повышать свою технологическую готовность к технологиям Индустрии 4.0 [5], а технологическая готовность требует в свою очередь развития использования возможностей информационно-коммуникационных технологий [6].

В условиях массовой цифровизации бизнеса встает задача эффективного использования данных о клиентах, которыми располагают компании. Формирующийся информационный рынок открывает предприятиям невиданный ранее доступ к данным о потребителях. Эти данные могут быть ценными по многим причинам, в частности потому, что они могут быть информативными в отношении запросов и готовности потребителя платить, таким образом, растущий доступ к потребительским данным ставит новые задачи по их использованию и регулированию.

Первый уровень использования данных о клиентах – это внедрение CRM-систем, которые позволяют сотрудникам компании вручную отслеживать существующие данные о клиентах и информацию об истории взаимодействия с ними.

Второй уровень использования данных о клиентах – автоматизация взаимодействия, когда CRM-система реализует сценарии взаимодействия [7]. В новом, быстро развивающемся мире предпочтения и вкусы клиентов быстро эволюционируют, создавая огромные проблемы для предприятий. Перемены в поведении клиентов приводят к быстрому устареванию продуктов и услуг, что требует пересмотра рабочего процесса и применения инновационных продуктов и сервисов для обеспечения потребностей и комфорта клиентов [3]. Соответственно, компаниям необходимо развивать технологические навыки и потенциалы для разработки динамических возможностей CRM, чтобы адаптироваться к столь быстро изменяющемуся деловому окружению [2].

Следует отметить, что система управления взаимоотношениями с клиентами объединяет и анализирует клиентские данные, полученные в результате формальных и случайных отношений между партнерами в сети, включая поставщиков услуг и клиентов [8]. Важная роль в этом случае отводится предиктивной аналитике для получения оптимальных решений, сведения к миниму-

му уровня неопределенности, своевременного реагирования на изменения в показателях различных видов бизнеса и точного управления рисками [9]. Базовым сценарием реализации такого взаимодействия является автоматизированная генерация персонализированных предложений [10]. При этом данный сценарий может быть реализован либо в виде системы поддержки сотрудника компании (в виде подсказок, какие предложения стоит сделать клиенту в процессе диалога), либо путем передачи предложений непосредственно клиенту. Подчеркивается, что способность управлять отношениями с клиентами с помощью искусственного интеллекта, что жизненно важно для конкурентоспособности, позволит компаниям понять меняющиеся предпочтения клиентов и добиться высокой результативности отношений с ними [11, 12]. Результатом оптимизации взаимодействия является устойчивая общественно значимая продуктивность [13, 14].

Цель исследования – разработка универсальной методики прогнозирования потребностей клиентов экосистемы широкого класса, в частности для возможности формирования персонализированных торговых предложений.

Материалы и методы исследования

Разработка и тестирование описываемой методики проводились на основе больших данных из сферы пассажирских авиаперевозок (PNR и ETLR записи). Используемые в разработке методики данные при этом включали как количественные, так и качественные показатели. К данным применялись различные известные подходы к кластеризации и анализировались полученные распределения на предмет их интерпретируемости. На основе проведенного анализа были учтены недостатки и достоинства применяемых подходов и разработан оригинальный метод кластеризации, обладающий большей гибкостью, возможностями контроля со стороны пользователя и меньшими требованиями к вычислительным мощностям.

Результаты исследования и их обсуждение

С развитием технологий накопление и обработка больших массивов данных стали 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 менее важной задачей является эффективная и надлежащая обработка собранных данных, которая должна обеспечить компанию информацией, соответствующей принятию решений и экономическим целям.

Такая обработка информации всегда не тривиальна, так как зависит от:

- задач организации;
- объема располагаемых данных;
- целостности данных;
- типа данных (качественные, количественные или смешанные);
- квалификации лиц, принимающих решения.

В итоге любой инструмент анализа данных требует существенной адаптации под каждую конкретную компанию, что вызывает сложности с созданием подобных инструментов анализа данных и ставит 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 новым, без истории взаимодействия с нашей компанией, так и пользователем, о котором у компании присутствует уже значительный массив исторических данных. Такая постановка задачи 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 объявлениям, покупки. Каждому из совершенных

пользователем действий с конкретным предложением-рекламой присваивался определенный балл: простой показ предложения считался за 1 балл, клик по предложению – 2 балла, а покупка – 10 баллов. Суммируя эти баллы по каждому из показов предложения данному клиенту, высчитывался средний рейтинг предложения дополнительной услуги у конкретного клиента.

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, а также 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) – 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.



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

Схема работы алгоритма кластеризации

Однако применение классического кластерного анализа (k-means, сети Кохонена и др.) не дало 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 [15] через присвоение клиентам ярлыков («тэгов», или «меток») на основе логических правил, например таких меток:

- #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.

В этом случае каждому пассажиру присваивается набор идентификаторов – тэгов, которые зависят от характеристик и поведения этого пассажира, что фактически является его «отпечатком» в нашей системе.

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.

Главным преимуществом предлагаемого подхода является его прозрачность для компании, в отличие от технологии искусственного интеллекта, являющейся 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 в нее новые уровни.

Заключение

В результате проведенного исследования на примере данных записей PNR и ETLR из сферы пассажирских авиаперевозок был разработан оригинальный подход к прогнозированию потребностей клиентов с применением методов кластеризации базы данных клиентов на основе тэгов. Данный метод был разработан на основе проведенного анализа преимуществ и недостатков классических подходов к кластеризации клиентских баз и расчета персональных рекомендаций. Используемые из записей PNR и ETLR данные включали в себя как качественные, так и количественные показатели, что делает разработанный метод универсальным и применимым абсолютно к любым типам данных о клиентах.

Разработанная методика основывается на пересекающейся кластеризации клиентской базы через присвоение клиентам тэгов на основе созданной пользователем системы логических правил. Имеющиеся же в информационной системе записи о совершенных клиентами действиях позволяют рассчитывать наиболее вероятное поведение похожих клиентов, таким образом прогнозируя их потребности.

Предложенный подход отличается высокой прозрачностью для пользователя, что

повышает доверие к системе, возможностью гибкого управления пользователями, поскольку коэффициентами и правилами маркировки можно управлять, а вычислительная мощность, необходимая для расчета рекомендаций, невелика.

Список литературы

1. 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. Т. 13, № 5. DOI: 10.18334/epp.13.5.117670.
2. Абрамов В.И., Абрамов И.В., Поливанов К.В., Семенков К.Ю. Цифровая трансформация системы управления отношениями с клиентами // Вопросы инновационной экономики. 2023. Т. 13, № 1. С. 289–306. DOI: 10.18334/vines.13.1.117051.
3. Dubey R., Gunasekaran A., Childe S.J., Bryde D.J., Giannakis M., Foropon C., Hazen B.T. Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism: A study of manufacturing organisations // International Journal of Production Economics. 2020. Vol. 226. Article 107599.
4. Lee H.L. The triple-A supply chain. Harvard Business Review. 2004. Vol. 82 (10). P. 102–113.
5. Samaranyake P., Ramanathan K., Laosirihongthong T. Implementing industry 4.0-A technological readiness perspective // In 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). 2017. P. 529–533.
6. Gupta M., George J.F. Toward the development of a big data analytics capability // Information & Management. 2016. Vol. 53(8). P. 1049–1064.
7. 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.
8. Bag S., Srivastava G., Bashir M.M.A., Kumari S., Giannakis M., Chowdhury A.H. Journey of customers in this digital era: Understanding the role of artificial intelligence technologies in user engagement and conversion // Benchmarking: An International Journal, ahead-of-print. 2021. DOI: 10.1108/BIJ-07-2021-0415.
9. Моськин А.В., Макаренко Е.А. Стратегическое развитие инновационных форм управления рисками // Стратегическое развитие социально-экономических систем в регионе: инновационный подход. 2020. С. 401–406.
10. 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. DOI: 10.35854/1998-1627-2023-3-335-344.
11. Herman L.E., Sulhaini S., Farida N. Electronic customer relationship management and company performance: Exploring the product innovativeness development // Journal of Relationship Marketing. 2021. V. 20 (1). P. 1–19.
12. Saura J.R., Ribeiro-Soriano D., Palacios-Marqués D. Setting B2B digital marketing in artificial intelligence-based CRMs: A review and directions for future research // Industrial Marketing Management. 2021. V. 98. P. 161–178.
13. Vesal M., Siahtiri V., O’Cass A. Strengthening B2B brands by signalling environmental sustainability and managing customer relationships // Industrial Marketing Management. 2021. Vol. 92. P. 321–331.
14. Edwin Cheng T.C., Kamble S.S., Belhadi A., Ndubis N.O., Lai K.H., Kharat M.G. (2021). Linkages between big data analytics, circular economy, sustainable supply chain flexibility, and sustainable performance in manufacturing firms // International Journal of Production Research. 2021. P. 1–15.
15. 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.