

УДК 004.8:007:658

## ТЕХНОЛОГИЯ КЛАСТЕРИЗАЦИИ И ГЕНЕРАЦИИ ПЕРСОНАЛИЗИРОВАННЫХ ТОРГОВЫХ ПРЕДЛОЖЕНИЙ ДЛЯ ПАССАЖИРСКИХ АВИАПЕРЕВОЗОК

<sup>1</sup>Гордеев В.В., <sup>2</sup>Столяров А.Д., <sup>2</sup>Абрамов В.И.

<sup>1</sup>Общество с ограниченной ответственностью «АЭРОЛАБС», Москва,  
e-mail: v.gordeev@aerolabs.aero;

<sup>2</sup>Национальный исследовательский ядерный университет МИФИ, Москва,  
e-mail: mr.alexst@gmail.com, viabramov@mephi.ru

Предметом статьи является раскрытие предпосылок и подхода к разработке технологии кластеризации пассажиров авиакомпании для дальнейшей генерации персональных предложений. В статье обсуждается важность создания решений, позволяющих упростить компаниям продажи дополнительных услуг – как собственных, так и партнерских, при этом повысив вероятность приобретения данных услуг клиентами за счет персонализации. Раскрываются типы используемой для этого информации о пассажирах и подходы к ее обработке – на основе искусственного интеллекта и на основе формальных правил. Приводится описание алгоритмов работы итогового продукта, обучаемого на основе действий пользователей и генерирующего персональные предложения для клиентов на основе проведенной кластеризации. На основе описываемых в статье подходов было разработано программное обеспечение, уже внедренное в хозяйственную деятельность и доказавшее справедливость и эффективность применяемых подходов. Технология, описываемая в статье, разработана специализированно для авиакомпаний, однако при незначительной модификации для учета отраслевой специфики может быть внедрена во множестве других сфер, таких как железнодорожные пассажирские перевозки, банковский сектор, страховая сфера, розничная торговля, медийная сфера, и в иных секторах хозяйственной деятельности.

**Ключевые слова:** рекомендательная система, кластеризация клиентов, сбор и обработка данных, продажи дополнительных услуг, программное обеспечение

## INNOVATIVE TECHNOLOGY FOR GENERATION OF PERSONALIZED OFFERS FOR AIR COMPANY PASSENGERS

<sup>1</sup>Gordeev V.V., <sup>2</sup>Stolyarov A.D., <sup>2</sup>Abramov V.I.

<sup>1</sup>Aerolabs LLC, Moscow, Russian Federation, e-mail: v.gordeev@aerolabs.aero;

<sup>2</sup>National Research Nuclear University MEPHI, Moscow,  
e-mail: mr.alexst@gmail.com, viabramov@mephi.ru

The subject of the article is the disclosure of the prerequisites and approach to the development of technology for clustering airline passengers for further generation of personal offers. The article discusses the importance of creating solutions that make it easier for companies to sell additional services – both their own and partners, while increasing the likelihood of customers purchasing these services through personalization. The types of passenger information used for this and approaches to its processing are disclosed – based on artificial intelligence and on the basis of formal rules. The algorithms of the final product are described, which is trained on the basis of user actions and generates personal offers for customers based on the clustering. On the basis of the approaches described in the article, software has been developed that has already been introduced into economic activity and has proven the validity and effectiveness of the approaches used. The technology described in the article was developed specifically for airlines, however, with minor modifications to take into account industry specifics, it can be implemented in many other areas, such as rail passenger transportation, banking, insurance, retail, media and other sectors. economic activity.

**Keywords:** recommendation system, clustering of clients, data collection and processing, sales of additional services, software

В условиях ВАНИ-мира (акроним от английских слов: «хрупкий», «тревожный», «нелинейный» и «непонятный»), сложившегося в результате развития Индустрии 4.0, и становления шестого технологического уклада, требуются новые подходы к управлению компаниями, основанные на активном использовании новых бизнес-моделей и инновационных цифровых технологий. Задача цифровой трансформации экономики и увеличения темпов экономического развития страны актуальна как никогда, поэтому требуются иные подходы к управлению с использованием инновационных цифровых технологий, дающие но-

вые способы наращивания эффективности работы предприятий. Важным условием и фактором успешного проведения цифровой трансформации является повышение цифровой зрелости, которое выражается в степени готовности предприятия к запланированным переменам [1]. Показано, что стратегическая задача цифровой трансформации компании заключается в построении практически жизнеспособного цифрового двойника, который будет описывать взаимосвязь между цифровыми активами и видами деятельности, моделируя взаимодействие между различными источниками данных в организации [2]. Важным факто-

ром при этом является использование предиктивной аналитики, которая объединяет в себе множество методов машинного обучения, статистики, моделирования, теории игр, на основе которых производится анализ исторических данных, составляются предсказания будущего состояния объекта и выявляются закономерности, а также определяются возможные риски и возможности. За счет интерпретации выявленных закономерностей предиктивная аналитика помогает создать персонализированный подход по доставке информации при работе с клиентами и предугадать их потребности и желания. Преимущества включения предиктивного анализа в жизненный цикл организации заключаются в предоставлении интеллектуального анализа для принятия оптимальных решений, минимизации неопределенности, в точном управлении рисками, своевременном реагировании на изменение различного рода показателей эффективности предприятия.

Целью исследования является обоснование актуальности и эффективности предлагаемой технологии кластеризации и генерации персонализированных торговых предложений для пассажирских авиаперевозок с учетом отраслевой специфики и принципов построения такой технологии.

#### Материал и методы исследования

Был выполнен обзор аналитических отчетов и статистических данных по сектору пассажирских авиаперевозок, особенно относящихся к области продажи дополнительных услуг. Для разработки технологии кластеризации были проведены проблемные интервью с представителями авиакомпаний, на основе лучших практик работы с большими данными выбирались оптимальные подходы к решению выявленных проблем. В разработке использовались нейросети (методы искусственного интеллекта) и машинное обучение, а также применялись автоматизированные подходы на основе формальных правил обработки данных и кластеризации.

#### Результаты исследования и их обсуждение

Необходимость подобных решений, скорректированных под рынок пассажирских авиаперевозок, обусловлена структурой доходности авиакомпаний и активным ростом данного рынка [3]. Да, годы пандемии существенно сказались на рынке авиаперевозок, но наблюдается его возврат к прежним показателям и в 2022 году прогнозируется выручка 782 млрд долларов,

и это не намного меньше максимума выручки 838 млрд долларов в 2019 году [4]. При этом в пике в 2019 году было перевезено более 4,5 трлн пассажиров [5]. Каждому из этих пассажиров в той или иной форме делались предложения дополнительных сервисов, но с достаточно низкой степенью персонализации. Для более глубокого понимания нашего потенциала в этом секторе следует рассмотреть статистику по доходам авиакомпаний именно от продажи дополнительных сервисов. При этом по отрасли в целом доходы от продажи дополнительных сервисов занимают существенную долю в выручке авиакомпаний (более 10%) при огромной маржинальности, не сопоставимой с маржинальностью основной деятельности – непосредственно услуг по авиаперевозке (рис. 1).

При такой среднеотраслевой статистике видно, что лидеры рынка по внедрению дополнительных сервисов получают от них более половины своей выручки, а в рамках пандемии эта доля у многих компаний еще и существенно возросла (табл. 1). Таким образом, большинство авиакомпаний, в том числе крупнейших, имеют огромный потенциал по росту выручки уже сейчас – за счет интеграции грамотной системы продаж дополнительных услуг – как собственных (багаж, выбор места, питание и т.д.), так и сторонних (отели, трансферы, страхование). И для этого нужны, прежде всего, информационные решения, обеспечивающие информационную интеграцию с поставщиками услуг и внутренними системами авиакомпаний.

Таблица 1

Процент выручки от продажи дополнительных сервисов в доле общей выручки у избранных авиакомпаний [7]

№	Компания	2021	2019
1	Wizz Air	56,0%	45,4%
2	Frontier	54,9%	43,6%
3	Spirit	54,3%	47,0%
4	Allegiant	51,3%	46,5%
5	Viva Aerobus	44,8%	45,0%
6	Ryanair Group	44,7%	34,5%
7	Volaris	42,9%	38,5%
8	GOL	33,0%	17,0%
9	easyJet	31,4%	21,6%
10	Pegasus	30,8%	26,4%

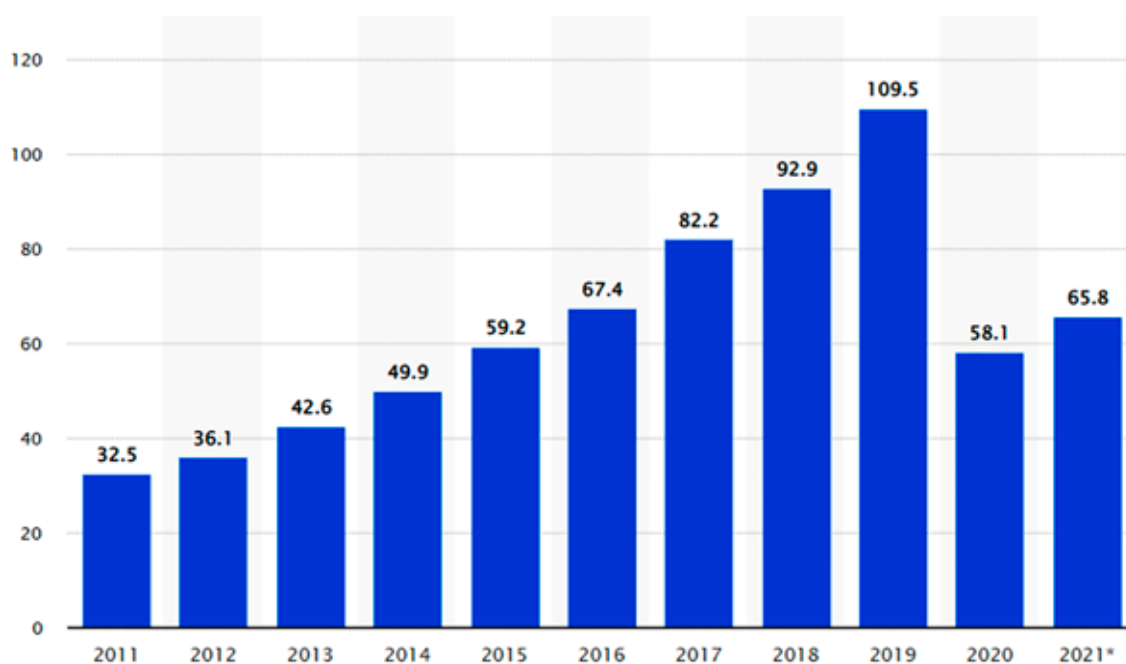


Рис. 1. Общая выручка от продажи дополнительных сервисов авиакомпаниями с 2011 по 2021 годы [6]

Такой продукт, стартовав в сфере пассажирских авиаперевозок, будет достаточно легко приспособлен и под иные сферы: железнодорожные компании; онлайн туристические агентства (и платформы по продаже билетов); туристические операторы; круизные компании; отельные агрегаторы (booking.com, expedia.com, kayak.com); веб-сервисы.

Объем данных в электронной коммерции систем авиакомпаний является одним из самых больших и может сравниться только с банковским сектором [8]. По сложности данных по каждому пассажиру и эволюции данных на каждом этапе полета системы электронной коммерции авиакомпаний также не уступают ни одной другой отрасли электронной коммерции. При этом решения по продажам на каждом этапе полета для пассажира должны приниматься в минимальное время. Все это фактически означает обработку массивных объемов данных, большое количество критериев при принятии решений, а также ограниченность времени принятия решений. Кроме того, постоянное добавление критериев, объектов продажи и обогащение данных о пассажире не позволяют задать четкие правила выбора времени, метода и объекта продажи: требуется постоянная коррекция принимаемых решений для максимизации дохода [9].

С точки зрения сбора и обработки информации интерес представляют три больших класса данных о пассажирах.

**1. Статические данные.** К ним относится информация о пассажире, которая не изменяется со временем и не накапливается. Ключевыми записями, относящимися к статическим данным, являются: паспортные данные; пол; возраст; дата рождения; знак зодиака; национальность; рост; страна рождения.

**2. Исторические данные.** К таким данным относятся транзакционные данные о пассажире, накапливаемые о пассажирах в процессе приобретения услуг компании и любом взаимодействии с ней: количество полетов; средний чек; история полетов; история покупок; компаньоны; время покупок; устройства, с которых делается заказ.

**3. Классификация пассажира:** данные (метки), присваиваемые пассажирам в процессе обработки записей о них системой кластеризации. Это могут быть как абстрактные кластеры, сформированные как с помощью обработки базы данных пассажиров методами машинного обучения и/или искусственного интеллекта, так и вручную – на основе присвоения пассажирам (клиентам) тех или иных меток на основе логических правил, заданных разработчиком (администратором/маркетологом), например: веган; любитель горнолыжного отдыха; часто летающий на Бали; летает с семьей; трое детей; летает по работе; VIP.



Рис. 2. Сбор статических и исторических данных из разных информационных систем

Сбор статических и исторических данных при этом является нетривиальной задачей, так как значимые записи содержатся в различных информационных системах авиакомпании, причем механизмы интеграции данных системы нестандартны и для каждого клиента свои. Чаще всего данные собираются из систем, изображенных на рисунке 2. К таким системам относятся следующие.

**Системы управления лояльностью пассажиров.** Как правило, основная функция данных систем – накопление подарочных миль и закрепление их за конкретным человеком. Из данного функционала вытекает критическая необходимость правильной и однозначной идентификации человека. Из-за этого из всех остальных информационных систем авиакомпании системы управления лояльностью содержат наиболее достоверные идентификационные данные пассажира.

**Системы бронирования (Reservation System).** Содержат данные об истории бронирований, выгрузка которых привязывается к пользовательским данным, полученным из систем управления лояльностью.

**Системы контроля отправки пассажиров (DCS – Departure Control Systems).** Необходимы для дополнительной верификации данных, так как содержат вручную проверенные в аэропорту при отправке пассажиров данные.

**Системы онлайн-регистрации на рейс.** Используются для получения верифици-

рованных контактных данных пассажиров, необходимых для дальнейшей отправки им сообщений и рекламных предложений.

**Данные колл-центров.** Наименее структурированные из всех и, как правило, требуют дополнительной обработки текстовых записей методами искусственного интеллекта. Однако полезны и в части получения дополнительных контактных данных клиентов.

Столь большое количество информационных систем для сбора данных обусловлено неполнотой данных о клиенте (пассажира), содержащейся в каждой из них по отдельности. Так, где-то из идентифицирующих клиента данных могут быть записаны только телефонный номер, или только почта, или только ID, или только паспортные данные, и т.д. Кроме того, часть этих данных может быть занесена неправильно. Происходить это может либо из-за особенностей структуры базы данных самой системы, либо из-за ошибок оператора при занесении персональных данных, либо из-за ошибок самого клиента. Из-за этого данные из различных систем приходится сливать и сводить в единую базу данных.

На основе описанных данных о пассажирах проводилась их кластеризация. Направлений кластеризации при этом было выбрано два.

1. Кластеризация методами искусственного интеллекта [10]. Позволяет автоматизировать обработку данных и за счет установленных критериев кластеризации получать некий универсально применимый

для маркетинговых целей результат. Очевидный плюс такого подхода – автоматизация. К отрицательным сторонам относится риск достаточно хаотичной кластеризации с точки зрения потребительских качеств, так как из-за специфики бизнес-процессов на вход системы подаются данные, весьма ограниченно связанные с портретом потребителя, прежде всего это статические данные, больше относящиеся к персональным. При этом из-за специфики авиаперевозок по каждому пассажиру имеется весьма небольшое количество транзакций, что затрудняет обучение алгоритмов и снижает релевантность кластеризации.

2. Кластеризация на основе формальных критериев. В данном случае каждому пользователю присваиваются «метки» на основе логических правил. Преимущества метода – однозначный контроль кластеризации со стороны авиакомпании, что повышает доверие к системе и возможность гибкой коррекции логики кластеризации.

Для внедрения был в итоге выбран именно подход, основанный на кластеризации клиентов с использованием формальных критериев с применением методов машинного обучения для формирования персональных предложений.

Алгоритм машинного обучения работает на основе вычисления метрики схожести всех клиентов компании, занесенных в базу данных. Для расчета метрики схожести используется множество тэгов, присваиваемых клиентам на основе установленных логических критериев. Например, тэг #familyTrip присваивается пассажирам, у которых больше 50% полетов – с семьей, тэг #dayOnly – пассажирам, летающим только днем, тэг #rainServicesOften – пассажирам, приобретающим дополнительные платные услуги. Таким образом, каждому пассажиру может быть присвоено множество тэгов. У каждого тэга при этом в настройках программного обеспечения администратором устанавливается вес – число от 1 до 10, выражающее субъективную важность для компании-клиента данного тэга-характеристики в расчете персональных предложений.

Для поиска похожих на него пассажиров происходит сопоставление целевого клиента со всеми остальными клиентами авиакомпании с целью выявления общих тэгов. Далее веса всех тэгов, общих для целевого и текущего клиента, складываются, в итоге получается абсолютный коэффициент сродства ( $K_c$ ):

$$K_{abs} = \sum \hat{w}_i,$$

где  $K_{abs}$  – коэффициент сродства;

$\hat{w}_i$  – вес каждого общего тэга у двух сравниваемых клиентов.

Для получения же удобной и наглядной относительной метрики сродства клиентов рассчитывался максимальный коэффициент сродства для целевого клиента  $K_{max}$ , равный сумме всех коэффициентов тэгов целевого пассажира:

$$K_{max} = \sum w_i,$$

где  $K_{max}$  – максимальный коэффициент целевого клиента;

$w_i$  – вес каждого из тэгов, присвоенного данному клиенту.

Тогда значение целевого для нас относительного коэффициента сродства целевого клиента с текущим будет выражаться следующим образом:

$$K_c = K_{abs} / K_{max}.$$

Данный коэффициент сродства далее применяется для расчета персональных предложений. В расчете используются данные трех сущностей:

- реестр сделанных клиентам предложений;
- реестр кликов клиентов по сделанным предложениям;
- реестр совершенных клиентами покупок.

Для первичного обучения программного обеспечения эти реестры либо заполняются предварительно сгенерированными данными, либо клиентам определенное время делаются случайные предложения и в дальнейшем фиксируется реакция на них.

Данные реестры фактически соответствуют разным этапам воронки продаж. При этом с точки зрения формирования предложений наиболее ценным для нас действием является совершенная клиентом покупка, менее ценным – клик по предложению. Для сообщения алгоритму машинного обучения данной ценности каждому из типов действия присваивается свой вес в баллах: простой показ предложения считается за 1 балл, клик по предложению – за 2 балла, а покупка – за 10 баллов.

Для примера, если Клиенту N только один раз делалось предложение, это было предложение такси класса «люкс» и клиент его в итоге заказал, то для предложения такси «люкс» у этого клиента будет  $1 + 2 + 10 = 13$  баллов. Если бы клиент данную услугу не стал приобретать, а только кликнул по ней, то рейтинг этого предложения у него составил бы лишь 3 балла. Если бы данному человеку на разных этапах взаимодействия с нашей компанией эта реклама была бы показана несколько раз, то в сумме набралось бы несколько просмотров, несколько кликов и несколько покупок и тогда суммарный

рейтинг данного предложения у этого клиента составил бы:

$$r_p = n_v \times k_v + n_c \times k_c + n_s \times k_s,$$

где  $r_p$  – рейтинг конкретного предложения у конкретного клиента (от  $p = proposal$ );

$n_v$  – количество показов данного типа предложения данному клиенту (от  $v = view$ );

$k_v$  – вес показа (по умолчанию равен 1);

$n_c$  – количество кликов по данному предложению данным клиентом (от  $c = click$ );

$k_c$  – вес клика (по умолчанию равен 2);

$n_s$  – количество покупок данной услуги данным клиентом (от  $s = sale$ );

$k_s$  – вес покупки (по умолчанию равен 10).

Итоговый же алгоритм расчета персональных предложений использует как метрику сродства клиентов, так и данные рейтинги предложений по каждому из клиентов. Алгоритм в текущей реализации модуля запускается непосредственно

в момент генерации предложения и состоит из следующих шагов (рис. 3).

- Для целевого клиента, которому делается предложение, подбирается ТОП-50 наиболее похожих клиентов из базы данных авиакомпании по метрике сродства, расчет которой описан выше.

- Для каждого из ТОП-50 похожих клиентов, а также для целевого клиента по всем когда-либо сделанным им предложениям рассчитываются рейтинги этих предложений.

- Каждый из рассчитанных для ТОП-50 рейтингов предложений умножается на коэффициент сродства клиента из ТОП-50 с текущим клиентом.

- Рейтинги предложений, рассчитанные для целевого клиента, умножаются на максимальный коэффициент сродства и дополнительно на 3 ( $k_i$  – вес целевого клиента) (вручную выставляемый коэффициент повышения релевантности).

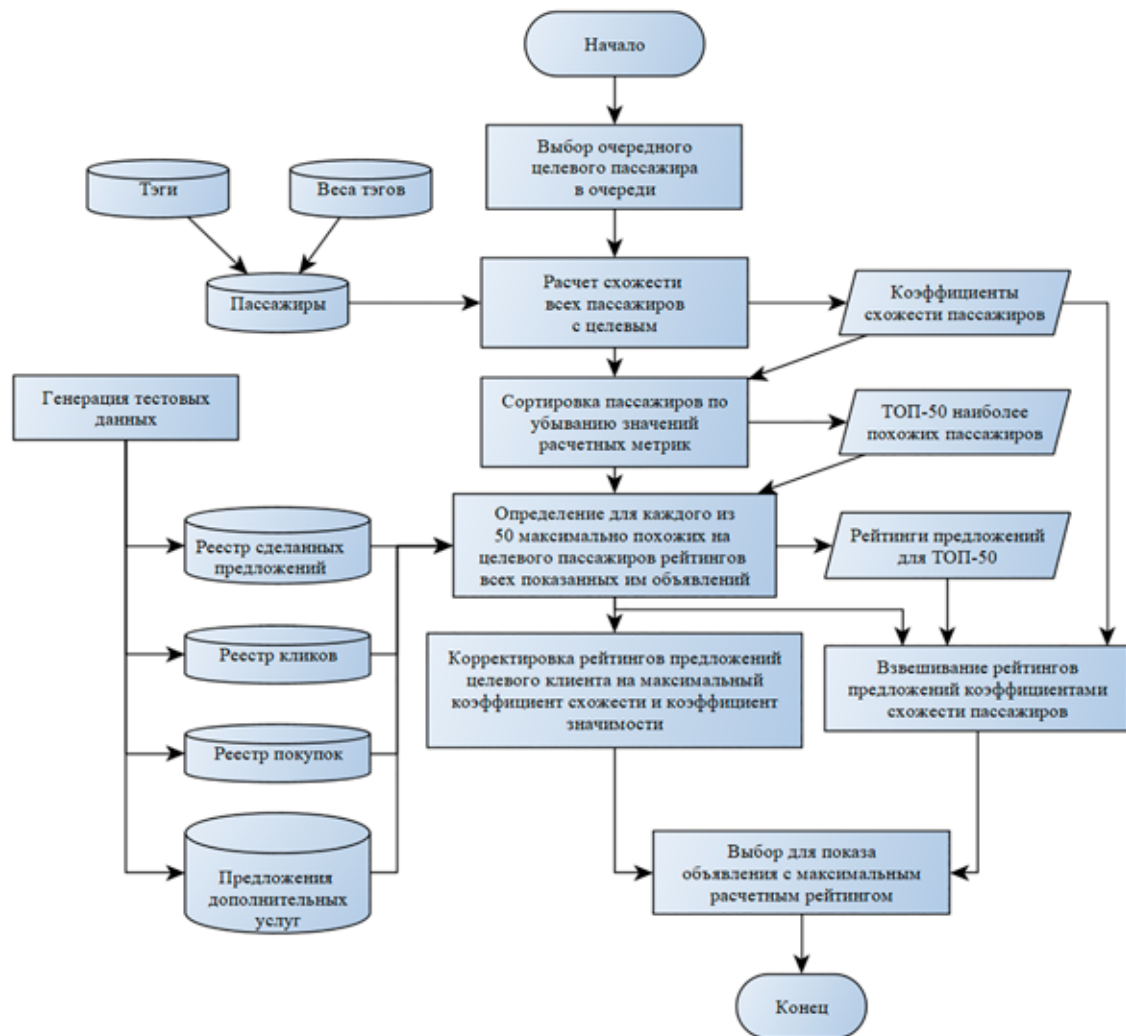


Рис. 3. Графическая схема работы алгоритма подбора персонального предложения

• По каждому из типов предложений суммируются все рассчитанные произведения рейтингов на коэффициенты средства.

• Целевому клиенту делается предложение, которое при таком расчете набрало наибольшее количество баллов.

Данный алгоритм поиска оптимального предложения можно описать следующей формулой:

$$MAX(R_i) = r_i^t \times K_{max} \times k_t + \sum_{j=1}^{50} r_i^j \times K_C^j,$$

где  $i$  – порядковый номер конкретного предложения;

$R_i$  – суммарный рейтинг конкретного предложения для данного клиента;

$r_i^t$  – рейтинг конкретного предложения из истории целевого клиента (от  $t = target$ );

$K_{max}$  – максимальный коэффициент средства целевого клиента;

$k_t$  – вес целевого клиента;

$j$  – порядковый номер клиента из ТОП-50 похожих на целевого;

$r_i^j$  – рейтинг конкретного предложения у  $j$ -го клиента из ТОП-50 похожих на целевого;

$K_C^j$  – коэффициент средства целевого клиента с  $j$ -тым клиентом из ТОП-50 похожих на целевого.

Или эту же формулу можно без каких-либо искажений использовать с абсолютным коэффициентом средства:

$$MAX(R_i) = r_i^t \times K_{max} \times k_t + \sum_{j=1}^{50} r_i^j \times K_{abs}^j$$

Если исходную формулу прописать в полностью развернутой форме, то модель поиска оптимального предложения будет выглядеть следующим образом:

$$\begin{aligned} MAX(R_i) = & \\ = & (n_{v_i}^t \times k_v + n_{c_i}^t \times k_c + n_{s_i}^t \times k_s) \times k_t \times \sum_{\alpha} w_{\alpha}^t + \\ + & \sum_{j=1}^{50} (n_{v_i}^j \times k_v + n_{c_i}^j \times k_c + n_{s_i}^j \times k_s) \times \frac{\sum_{\beta_j} \hat{w}_{\beta_j}^j}{\sum_{\alpha} w_{\alpha}^t}, \end{aligned}$$

где  $i$  – порядковый номер конкретного предложения;

$n_{v_i}^t, n_{c_i}^t, n_{s_i}^t$  – количество взаимодействий, совершенных целевым пользователем с  $i$ -м предложением, соответственно показов, кликов и покупок;

$k_v, k_c, k_s$  – веса соответствующих взаимодействий с рекламой (единые для всех клиентов), соответственно для показов, кликов и покупок;

$k_t$  – значимость показателей целевого клиента (по умолчанию 3);

$\alpha$  – суммарное количество тэгов у целевого клиента;

$w_{\alpha}^t$  – вес конкретного тэга у целевого клиента;

$j$  – порядковый номер клиента из ТОП-50 похожих на целевого;

$n_{v_i}^j, n_{c_i}^j, n_{s_i}^j$  – количество взаимодействий, совершенных  $j$ -м пользователем с  $i$ -м предложением, соответственно показов, кликов и покупок;

$\beta_j$  – количество общих тэгов у  $j$ -го клиента из ТОП-50 с целевым клиентом;

$\hat{w}_{\beta_j}^j$  – вес каждого их тэгов, общих для целевого клиента и  $j$ -го клиента из ТОП-50.

В результате разработанный метод дает возможность достаточно гибко управлять работой программным обеспечением со стороны клиента, не требуя при этом значительных вычислительных мощностей при проведении пересчета всех метрик в фоновом режиме по расписанию.

### Заключение

Методы искусственного интеллекта являются эффективным средством решения поставленных задач по кластеризации и генерации персональных предложений.

В соответствии с этим был разработан подход, позволяющий выполнять кластеризацию клиентов с использованием перечня тэгов, присваиваемых каждому клиенту на основе логических правил. Данные правила (тэги) гибко настраиваются, адаптируясь под пожелания заказчика при внедрении программного обеспечения.

Обучение программного обеспечения и генерация персональных предложений осуществляются путем расчета метрики схожести, которая учитывает присвоенные каждому клиенту тэги, сделанные ранее каждому клиенту предложения и действия клиентов, совершенные с данными предложениями.

Обучение программного обеспечения может проводиться как на основе генерации предварительных данных о действиях пользователей с объявлениями, так и на основе формирования предложений случайным образом в течение ограниченного периода времени по выбору заказчика.

Разработанный подход к генерации персональных торговых предложений позволяет гибко управлять процессами кластеризации клиентов и формирования предложений, обеспечивает простоту внедрения, исключая при этом скачкообразные нагрузки на серверную архитектуру, так как основные вычислительные операции

выполняются не по запросу пользователя, а в фоне по расписанию.

#### Список литературы

1. Абрамов В.И., Борзов А.В., Семенов К.Ю. Оценка готовности малых и средних предприятий к цифровой трансформации // Вопросы инновационной экономики. 2022. Т. 12. № 3. С. 1573-1596. DOI: 10.18334/vinec.12.3.115000.
2. Абрамов В.И., Бобоев Д.С., Гильманов Т.Д., Семенов К.Ю. Теоретические и практические аспекты создания цифрового двойника компании // Вопросы инновационной экономики. 2022. Т. 12. № 2. С. 967-980. DOI: 10.18334/vinec.12.2.114890.
3. Аладьев А.А. Перспективы развития российского рынка пассажирских авиаперевозок в среднесрочной перспективе 2021-2023 годов (прогнозная модель и предложения для авиакомпаний) // Modern Economy Success. 2021. № 1. С. 147-152.
4. Revenue of airlines worldwide 2003-2022 [Электронный ресурс]. URL: <https://www.statista.com/statistics/278372/revenue-of-commercial-airlines-worldwide/> (дата обращения: 12.02.2023).
5. Air transport, passengers carried [Электронный ресурс]. URL: <https://data.worldbank.org/indicator/IS.AIR.PSGR?end=2020&start=1998> (дата обращения: 12.02.2023).
6. Airline industry – ancillary revenue 2011-2021 [Электронный ресурс]. URL: <https://www.statista.com/statistics/788849/airline-industry-ancillary-revenue/> (дата обращения: 12.02.2023).
7. Sorensen J. The 2022 CarTrawler Yearbook of Ancillary Revenue: Годовой. IdeaWorksCompany.com LLC.
8. Гарифуллина Н.А. Большие данные в авиакомпании "Аэрофлот" // Свободный полет-2018. Задачи обработки больших данных в авиации: материалы III Всероссийской научно-практической конференции / Под общей редакцией С.С. Валеева, А.Д. Кулакова, И.А. Копылова (Жуковский-Уфа, 29–31 мая 2018 года). Уфа: Уфимский государственный авиационный технический университет, 2018. С. 46-49.
9. Абрамов В.И., Михайлов Д.М., Столяров А.Д. Эко-системный подход создания омниканальных коммуникаций и цифровых профилей клиентов // Управление бизнесом в цифровой экономике: сборник тезисов выступлений Четвертой международной конференции / Под общей редакцией И.А. Аренкова, М.К. Ценжарик. СПб.: Санкт-Петербургский государственный университет промышленных технологий и дизайна, 2021. С. 180-186.
10. Ahmed M., Seraj R., Islam S.M.S. The k-means Algorithm: A Comprehensive Survey and Performance Evaluation: 8 // Electronics. Multidisciplinary Digital Publishing Institute. 2020. Vol. 9, Is. 8. P. 1295.