

УДК 339.138:004.65

УПРАВЛЕНИЕ ВЗАИМОДЕЙСТВИЕМ С КЛИЕНТАМИ КОММЕРЧЕСКОЙ ОРГАНИЗАЦИИ НА ОСНОВЕ МЕТОДОВ СЕГМЕНТАЦИИ И КЛАСТЕРИЗАЦИИ КЛИЕНТСКОЙ БАЗЫ

А. Ф. Галямов¹, С. В. Тархов²

¹galyamov.airat@mail.ru, ²tarkhov@inbox.ru

ФГБОУ ВПО «Уфимский государственный авиационный технический университет» (УГАТУ)

Поступила в редакцию 17 апреля 2014 г.

Аннотация. Описываются методы поддержки принятия решений, используемые на начальных этапах процесса управления взаимодействием с клиентами коммерческих организаций. Рассматривается два класса методов разделения клиентской базы на группы: сегментация и кластеризация. Описаны методы ABC-анализа, XYZ-анализа, RFM-анализа и NPS-анализа, входящие в множество методов сегментации. В качестве механизма кластеризации клиентской базы рассмотрена нейронная сеть Кохонена. Показано практическое применение и результаты использования методов сегментации и кластеризации клиентской базы для управления взаимодействием с клиентами коммерческой организации.

Ключевые слова: управление взаимодействием с клиентами; CRM-система; клиентская база; сегментация; ABC-анализ; XYZ-анализ; RFM-анализ; NPS-анализ; кластеризация; сеть Кохонена.

Взаимодействие с клиентами является одной из ключевых разновидностей процессов, протекающих в социально-экономических системах. Клиентская база является одним из важнейших ресурсов коммерческих организаций и одновременно является сложным объектом управления, требующим применения специальных управленческих методов и средств. Повседневная деятельность любой как коммерческой, так и некоммерческой организации, предприятия, учреждения неотъемлемо связана с процессом взаимодействия с клиентами. Более того, в результате взаимодействия организации с клиентами заключаются сделки, осуществляются продажи товаров и оказание услуг, что является основным источником дохода коммерческих организаций. Как известно, целью функционирования любой коммерческой организации является получение максимального объема прибыли при существующих условиях-ограничениях (законодательных, социальных, нравственно-этических и др.).

Для достижения поставленных целей организации необходимо управлять процессом взаимодействия со своими клиентами. Данный процесс в большинстве случаев достаточно сложен и включает целый ряд этапов. К тому же, клиентская база организации, как правило, достаточно разнородна – разные клиенты обла-

дают различными характеристиками, им свойственны различные модели поведения. Управление взаимодействием с клиентами включает решение множества задач, таких как привлечение подходящих клиентов, их удержание (предотвращение оттока), классификация и приоритизация по различным критериям, рациональный выбор комплекса проводимых мероприятий, используемых способов взаимодействия для адресных обращений к клиентам для максимального удовлетворения их потребностей с целью выстраивания долгосрочных взаимовыгодных отношений с клиентами и максимизации прибыли от этого процесса. На начальных стадиях управления взаимодействием необходимо разделить клиентскую базу на группы клиентов со схожими характеристиками для повышения адекватности маркетинговых, торговых и сбытовых мер, предпринимаемых на последующих этапах процесса управления взаимодействием с клиентами.

В данной статье рассматриваются два класса методов разделения клиентской базы на группы: сегментация и кластеризация. При использовании методов сегментации заранее известно итоговое количество групп, а также правила принадлежности клиента к определенной группе. Напротив, количество групп-кластеров, получаемых в процессе кластеризации, заранее неиз-

вестно, группы формируются динамически в процессе работы метода. В статье производится формальная постановка задач сегментации и кластеризации клиентской базы, а также приводится описание сегментационных методов ABC-анализа, XYZ-анализа, RFM-анализа, NPS-анализа и метода кластеризации клиентской базы с использованием нейронной сети Кохонена.

ПОСТАНОВКА ЗАДАЧИ РАЗДЕЛЕНИЯ КЛИЕНТСКОЙ БАЗЫ НА ГРУППЫ

Очевидно, что разные клиенты обладают различными характеристиками, моделью поведения и, как следствие, имеют различный уровень текущей ценности для организации. Для эффективного и рационального управления взаимодействием с клиентами организации необходимо иметь возможность разделения клиентской базы на группы, выполняемого по различным критериям, на основе которого реализуются последующие этапы процесса управления взаимодействием с клиентами.

Разделение клиентской базы осуществляется на основе анализа характеристик различной природы (социально-демографических данных, сведений о фактах взаимодействия клиента с организацией – его поведенческой истории и т.п.) [1]. Пусть m – число рассматриваемых характеристик (атрибутов, признаков) клиентов, значимых для принятия решений при управлении взаимодействием с ними. Описание i -го клиента в теоретико-множественной форме представляет собой вектор:

$$X_i = (x_{i1}, x_{i2}, \dots, x_{im})^T, \quad (1)$$

где x_{ij} – значение j -й характеристики поведенческой истории i -го клиента; m – количество характеристик клиентов (размерность пространства состояний клиента), на основе которых проводится их разделение.

Под сегментацией в данной работе понимается процесс разделения клиентской базы на несколько сегментов-групп, базирующегося на правилах принадлежности клиента к определенному сегменту, оперирующим значениями характеристик клиентов x_{ij} . Сегментация клиентской базы проводится по одному или нескольким критериям. Обозначим множество критериев $C = \{C_k\}$. Клиентская база разделяется на сегменты по каждому критерию C_k , которому соответствует свое множество сегментов $S_k = \{S_k^l\}$. Для осуществления сегментации необходимо задать правила принадлежности клиента X_i с некоторыми характеристиками x_{ij} к одному из сегментов S_k^l по критерию C_k :

$$\text{для } \forall S_k^l \exists \{C_k, \text{Rule}(X_i, C_k, S_k^l)\}. \quad (2)$$

Вербально эта зависимость формулируется следующим образом: для любого сегмента S_k^l существует критерий C_k , а также правило принадлежности, оперирующее значениями характеристик клиента X_i .

Во множестве случаев правило $\text{Rule}(X_i, C_k, S_k^l)$ содержит границы интервала значений характеристики $x_{ij} - (x_{ij}^{\min}, x_{ij}^{\max})$, а условием принадлежности i -го клиента к сегменту S_k^l является попадание значения его характеристики x_{ij} в этот интервал:

$$\text{если } (x_{ij}^{\min} \leq x_{ij} \leq x_{ij}^{\max}) \rightarrow X_i \in S_k^l. \quad (3)$$

Следует отметить, что в формуле (3) показано правило принадлежности клиента к сегменту, оперирующее одним критерием и значением одной характеристики клиента. В том случае, если разделение клиентской базы производится по нескольким критериям, правило принадлежности будет содержать логические операторы «И» и «ИЛИ», а также оперировать значениями нескольких характеристик клиента.

ОБЗОР МЕТОДОВ СЕГМЕНТАЦИИ КЛИЕНТСКОЙ БАЗЫ

В данной статье рассматривается ряд статистических методов сегментации клиентской базы, общие сведения о которых приведены в табл. 1.

Столбец $N_{\text{Кр}}$ содержит количество критериев, по которым происходит разделение клиентской базы, а столбец $N_{\text{Сег}}$ – количество сегментов, на которые она разделяется.

ABC-анализ является обобщением эмпирического принципа Парето, согласно которому, в данном контексте, приблизительно 20 % клиентов приносят организации 80 % прибыли. В процессе проведения ABC-анализа производится ранжирование всего списка клиентов по одному или нескольким критериям с последующим разделением списка на три категории-группы. При этом существует несколько методов выделения групп: эмпирический, дифференциальный, метод суммы, многоугольника, касательных и др. В результате проведения ABC-анализа производится сегментация клиентов по критерию степени вклада в общий результат – доход организации от взаимодействия с клиентами (продаж и обслуживания). Для проведения ABC-анализа необходимо определить цель исследования, выбрать объекты и их значимые для анализа характеристики, сформировать показатель

степени вклада объектов, определить период времени для анализа, составить ранжированный список объектов по убыванию значения сформированного показателя и разбить полученный список на несколько сегментов. Целью проведения ABC-анализа клиентской базы является выявление наиболее ценных для организации клиентов для более рационального использования ресурсов, затрачиваемых на взаимодействие с ними.

XYZ-анализ используется для оценки характера потребления клиентами продуктов (товаров и услуг). В качестве показателя, характеризующего отдельного клиента, используется частота закупок (сделок). Разделение на категории производится на основе накопленной в базе данных информации о сделках, объемах прибыли и показателя вариации. В результате вычисления показателей, характеризующих характер потребления клиентами продуктов, формируется ранжированный список, упорядоченный по возрастанию среднеквадратического отклонения и по убыванию показателя вариации, рассчитываемого по формуле:

$$v = \frac{\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}}{\bar{x}}, \quad (4)$$

где x_i – количество сделок, осуществленных клиентом за i -й период (месяц, квартал и т.п.);

n – количество периодов; \bar{x} – среднее количество сделок за все n периодов.

Список клиентов, ранжированный по значению показателя вариации, делится на три класса: клиенты класса X осуществляют стабильные сделки с организацией, клиенты класса Y совершают нерегулярные периодические покупки, а клиенты класса Z совершают редкие эпизодические покупки.

Интегрированный ABC-XYZ-анализ позволяет получить двумерную совокупную матрицу сегментов размерностью 3×3 для распределения клиентской базы на девять сегментов в зависимости от доли клиентов в доходах организации и от стабильности осуществляемых ими сделок одновременно. Для каждого сегмента S_{ij} определены два диапазона ($V_{\min}^i \dots V_{\max}^i$) и ($F_{\min}^j \dots F_{\max}^j$), определяющих границы сегментов по критерию ценности и частоты сделок соответственно. Правило попадания клиента в сегмент S_{ij} выражается логическим правилом:

$$\begin{aligned} &\text{если } (V_{\min}^i \leq V_{X_m} \leq V_{\max}^i) \cap \\ &\cap (F_{\min}^j \leq F_{X_m} \leq F_{\max}^j), \quad (5) \\ &\text{то } X_m \in S_{ij}, \end{aligned}$$

где X_m – m -й клиент организации (множество значений его характеристик); V_{km} – значение комплексного показателя ценности клиента; F_{km} – значение частоты осуществления сделок m -м клиентом организации.

Таблица 1

Классификация методов сегментации клиентской базы

Название	$N_{кр}$	$N_{сер}$	Цель	Критерии
ABC-анализ	1	3	Выявление клиентов, вносящих наибольший вклад в результаты процесса взаимодействия с организацией.	Доход от клиента / прибыль / объем продаж.
XYZ-анализ	1	3	Выявление постоянных клиентов, стабильно осуществляющих сделки.	Коэффициент вариации частоты сделок.
RF-анализ	2	≤ 25	Выявление наиболее активных клиентов, склонных к повторному осуществлению сделок.	Давность, частота и стоимость сделок.
RFM-анализ	3	≤ 125	то же	то же
ABC-XYZ-анализ	2	9	Выявление клиентов, стабильно осуществляющих сделки на большие суммы.	Доход от клиента, частота сделок.
NPS-анализ	1	3	Выявление активных клиентов – сторонников организации.	Вероятность рекомендации.

RFM-анализ базируется на получении значений таких частных показателей текущей ценности клиента для организации, как давность осуществления последней сделки R (*Recency*) и частота осуществления сделок F (*Frequency*) в течение периода для всех клиентов. Следует отметить, что для упрощения процедуры сегментации и повышения ее наглядности (для представления результатов сегментации в виде точечной диаграммы на плоскости) показателем суммы денег, затраченных клиентом на осуществление сделок за период (*Monetary*), пренебрегают по причине его сильной корреляции с показателем частоты осуществления сделок. Сегментация на основе показателей R и F включает несколько этапов:

1. Все множество значений показателя давности R делится на пять интервалов. При этом в пятый интервал попадают наименьшие значения давности (показатели клиентов, недавно осуществивших сделки), а в первый – наибольшие.

2. Множество значений показателя частоты F также делится на пять интервалов. В пятый интервал попадают наибольшие значения частот (показатели клиентов, осуществлявших наибольшее количество сделок за период), а в первый – наименьшие значения частот.

3. Строится матрица (сетка) для разделения клиентов на 25 сегментов на основе пар значений показателей R и F . Наилучшим клиентом в контексте данной сегментации является клиент со значением $R = 5$ и $F = 5$.

4. Для удобства использования на практике 25 RF -сегментов объединяются в большие сегменты по определенным правилам, например, на основе расчета коэффициентов перехода для каждого сегмента.

Полученные данные о принадлежности клиентов к определенным сегментам служат базовым критерием в процессе поддержки принятия решений о проведении комплексов мероприятий и выбора рациональных способов взаимодействия (маркетинговых каналов) [2].

NPS-анализ используется для разделения клиентской базы на основе оценки лояльности клиентов. Индекс NPS (*Net Promoter Score*) является показателем лояльности (приверженности) клиентов продукту, бренду или организации, характеризует готовность клиента рекомендовать продукт своему окружению и используется для оценки готовности к повторному осуществлению сделок. Процесс

измерения показателя лояльности клиентов включает ряд этапов:

1. Всем клиентам (при их небольшом количестве) или участникам фокус-группы предлагается ответить на вопрос «Какова вероятность того, что вы рекомендуете приобрести продукт или взаимодействовать с организацией своим коллегам, друзьям и знакомым?» по десятибалльной шкале. При этом ответ «ноль» соответствует ответу «ни в коем случае не порекомендую», а ответ «десять» – «рекомендую точно».

2. На основе полученных значений оценок клиентская база или фокус-группа разделяется на несколько сегментов, например, на три группы: клиенты с оценкой девять-десять баллов – сторонники продукта или организации, с оценкой семь-восемь баллов – нейтральные клиенты, с оценкой шесть баллов и ниже – критики.

3. Итоговый индекс NPS рассчитывается как разница количества сторонников S и критиков C продукта или организации: $NPS = S - C$.

Хорошим считается значение NPS на уровне 5–10 %.

МЕТОД КЛАСТЕРИЗАЦИИ КЛИЕНТСКОЙ БАЗЫ НА ОСНОВЕ СЕТИ КОХОНЕНА

Целью проведения кластерного анализа клиентской базы является ее разделение на группы клиентов со схожими признаками, основанное на оценке значений множества характеристик клиентов (их поведенческой истории). В отличие от процесса сегментации, при проведении кластеризации заранее не известно количество формируемых кластеров-групп клиентов, также как и правила принадлежности клиента к определенному кластеру. По этой причине кластерный анализ относится к классу задач машинного обучения (*machine learning*). Кроме того, кластеризация является одной из задач класса *Data Mining* и позволяет обнаруживать в наборе «сырых» необработанных данных поведенческой истории клиентов скрытые закономерности с целью проведения адресной персонализированной маркетинговой политики в отношении клиентов.

Существует большое количество систем классификации методов кластеризации и еще больше самих методов, среди которых выделяются наиболее популярные методы K -средних и K -медиан, G -средних, а также методы, основан-

ные на применении нейронных сетей и карт Кохонена, которые рассмотрены в данной статье.

Нейронная сеть Кохонена, используемая для кластеризации множества клиентов, содержит два слоя нейронов, ее схема показана на рис. 1.

Количество нейронов, составляющих входной слой сети, соответствует количеству рассматриваемых характеристик m . Входные нейроны не принимают участие в процессе обучения сети, они только передают значения характеристик клиентов из исходной выборки нейронам выходного слоя.

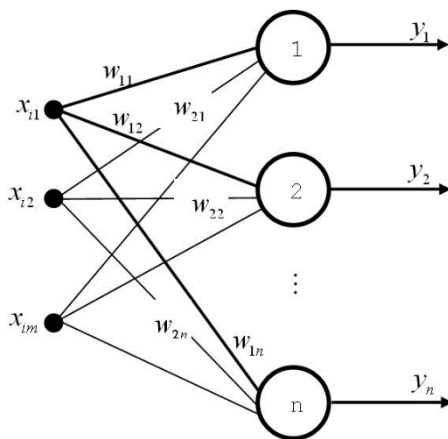


Рис. 1. Структура нейронной сети Кохонена, используемой для кластеризации клиентов по значениям их характеристик

Количество нейронов выходного слоя сети должно соответствовать числу кластеров, которые будут сформированы. Как правило, эта величина подбирается экспериментальным путем. Увеличение числа нейронов выходного слоя в ряде случаев позволяет добиться повышения детализации результатов кластерного анализа. Веса j -го нейрона выходного слоя представляют собой вектор:

$$W_j = (w_{1j}, w_{2j}, \dots, w_{mj})^T, \quad (6)$$

где w_{mi} – значение веса m -й связи j -го нейрона выходного слоя сети.

Перед непосредственным применением сети Кохонена для кластеризации клиентов требуется провести ее обучение, которое заключается в подстройке весов нейронов в процессе подачи на вход сети значений характеристик клиентов из обучающего множества.

Обучение сети Кохонена представляет собой итеративный процесс, включающий несколько этапов:

1. Установка начальных значений параметров сети: начальной скорости обучения η , радиуса обучения R . Генерация начальных

значений весов нейронов выходного слоя случайным образом в диапазоне от 0 до 1.

2. Возбуждение сети путем подачи на ее вход нормированных значений характеристик очередного клиента из множества обучающей выборки.

3. Для каждого нейрона выходного слоя вычисляется метрика расстояния между векторами характеристик клиента и весов нейрона. Наиболее распространенной мерой близости векторов является евклидово расстояние. В этом случае расстояние между вектором значений характеристик n -го клиента из обучающей выборки и вектором весов j -го нейрона вычисляется по формуле:

$$d_j = d(W_j, X_n) = \sqrt{\sum_i (w_{ij} - x_{in})^2}, \quad (7)$$

4. Определяется номер нейрона j_{\min} выходного слоя сети, значение расстояния d_j для которого является минимальным, согласно формуле:

$$j_{\min} = \arg \min_j (d_j). \quad (8)$$

Данный нейрон считается «победителем». Согласно заданному радиусу обучения R определяются все нейроны, для которых будет произведена подстройка их весов [3].

5. Для всех нейронов, находящихся в радиусе обучения R от нейрона-«победителя», производится корректировка (подстройка) значений их весов по следующей формуле:

$$w_{ij, \text{new}} = w_{ij, \text{cur}} + \eta(x_{ni} - w_{ij, \text{cur}}), \quad (9)$$

где $w_{ij, \text{new}}$ – новое значение веса i -й связи j -го нейрона; $w_{ij, \text{cur}}$ – текущее значение веса; η – коэффициент скорости обучения, принимающий значение в интервале от 0 до 1.

6. Изменение радиуса обучения и коэффициента скорости обучения согласно заданным параметрам алгоритма.

7. Если в множестве обучающей выборки остались необработанные данные клиентов, то осуществляется переход к пункту 2, иначе – к пункту 8.

8. Оценка значения критерия остановки обучения, прекращение или продолжение обучения в зависимости от этого значения. В качестве критериев остановки обучения используются такие условия, как ограничение количества циклов обучения, стабилизация выхода сети (объекты кластеризации не

переходят из одного кластера в другой), незначительность изменения весов нейронов.

Следует отметить, что при использовании обученной сети Кохонена каждому нейрону выходного ее слоя ставится в соответствие свой кластер. Зачастую одному нейрону ставится в соответствие один кластер. Применение обученной нейронной сети для кластеризации клиентов на основе данных их поведенческой истории включает ряд этапов:

1. Подача на вход сети вектора, содержащего нормированные значения характеристик поведенческой истории клиента.

2. Вычисление значений расстояний d_j между вектором входного воздействия X_i и векторами весов W_j всех нейронов выходного слоя сети по формуле (6). Нахождение номера нейрона j_{\min} по формуле (7), расстояние между вектором весов которого и вектором входного воздействия является минимальным.

3. Вектор входного воздействия относится к кластеру, соответствующему нейрону с номером j_{\min} .

4. Как правило, процесс кластеризации и интерпретация его результатов носят субъективный характер. В большинстве случаев процедуру кластерного анализа клиентской базы целесообразно проводить несколько раз, изменяя структуру и параметры нейронной сети Кохонена – количество нейронов выходного слоя, радиус обучения, начальное значение скорости обучения и правило его изменения в течение этого процесса. Результаты кластеризации клиентов, полученные при различных конфигурациях сети, оцениваются согласно заданному критерию качества с целью выбора лучшего результата. В роли критерия оценки качества результатов кластеризации клиентской базы, наряду с другими величинами, используется следующая функция:

$$Q = \frac{d_{\text{int}}}{d_{\text{ext}}}, \quad (10)$$

где d_{int} – среднее расстояние между векторами характеристик клиентов, попавших в один кластер (внутрикластерное расстояние); d_{ext} – среднее расстояние между векторами характеристик клиентов, попавших в различные кластеры (межкластерное расстояние).

Также следует отметить, что каждому полученному кластеру необходимо дать содержательное описание.

РЕЗУЛЬТАТЫ ПРАКТИЧЕСКОЙ РЕАЛИЗАЦИИ МЕТОДОВ СЕГМЕНТАЦИИ И КЛАСТЕРИЗАЦИИ КЛИЕНТСКОЙ БАЗЫ

Для иллюстрации работы методов сегментации была проанализирована клиентская база промышленного предприятия, содержащая сведения об осуществленных сделках за 2012 г. Количество активных клиентов за этот период составило 529, ими было осуществлено более 9,9 тыс. сделок на общую сумму 6,1 млрд руб., средняя сумма сделки составила порядка 615 тыс. руб.

Результаты проведенного анализа показали, что к сегменту А относится 106 клиентов (20,1 % от общего количества), приносящих промышленной организации совокупный доход 80 % от общей суммы. К сегменту В относится 138 клиентов (26,2 % клиентской базы), приносящих организации еще 15 % от ее дохода. В сегмент С вошли 283 клиента (53,7 % от общего числа), приносящие организации 5 % ее дохода. На рис. 2 представлена диаграмма, построенная по результатам ABC-анализа, на ней показаны кумулятивная кривая (накопительный процент) и убывание сумм сделок клиентов. По оси X отложено количество клиентов, по оси Y – сумма сделок отдельных клиентов и накопительный процент общей суммы. На диаграмме показано, что 20,1 % клиентов совершили сделки объемом 80 % от суммарного объема сделок предприятия.

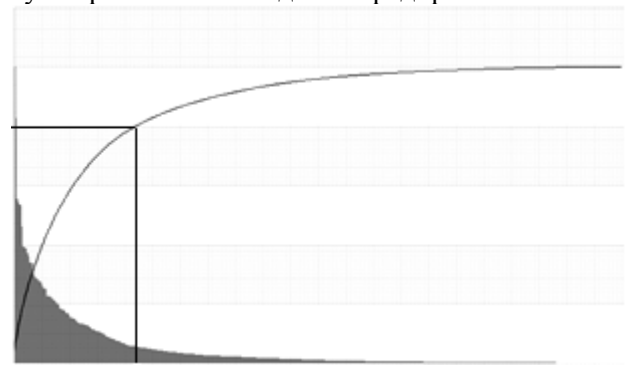


Рис. 2. Диаграмма результатов ABC-анализа

Анализ поведенческой истории клиентов Предприятия по критериям давности осуществления сделок R и их частоты F (RFM-анализ) позволил провести сегментацию клиентской базы и выявить клиентов, наиболее склонных к повторным сделкам за счет действия принципа поведенческой инерции. Минимальное значение частоты осуществления сделок F_{\min} равно 0,003, максимальное F_{\max} равно 0,814. Максимальная давность сделки R_{\max} составила 364 дня, минимальная R_{\min} –

3 дня. Множество значений давности R и частоты F разделено на пять равных интервалов, размер сегмента (шаг) F_{step} равен 0,1622, размер сегмента $R_{\text{step}} = 72,2$ дня. Из распределения количества клиентов по различным сегментам RF -матрицы видно, что наиболее многочисленным является сегмент клиентов, осуществлявших сделки сравнительно недавно (в течение 75 дней), но с небольшой частотой в течение года (от 0,003 до 0,165 сделок в год).

Для разделения клиентов на сегменты с использованием метода XYZ-анализа были использованы следующие граничные значения коэффициента вариации количества сделок по кварталам: к классу X относятся клиенты с коэффициентом вариации от 0 до 0,5; к сегменту Y – со значением этого коэффициента от 0,5 до 3,0; к классу Z – клиенты с коэффициентом вариации свыше 3,0. В результате проведенного XYZ-анализа выявилось, что сегмент X включает 219 клиентов, сегмент Y – 156 клиентов, а сегмент Z – 114 клиентов. Обладая знаниями о принадлежности клиента к одному из указанных классов, лицо, принимающее решения, имеет возможность прогнозировать объем сделок, а также предсказывать поведение клиентов – количество совершаемых ими сделок за определенный период.

В качестве иллюстрации процесса обучения нейронной сети Кохонена и кластерного анализа клиентской базы на ее основе рассмотрен следующий пример: база данных поведенческой истории клиентов содержит сведения о деятельности 105 клиентов промышленного предприятия за 2013 г. Для простоты и наглядности эксперимента кластеризация проводилась по двум параметрам – частоте осуществления сделок и давности осуществления последней сделки (в днях).

Графики изменения значений среднего внутрикластерного, среднего межкластерного расстояний и значения показателя качества кластеризации в зависимости от количества нейронов выходного слоя сети Кохонена показан на рис. 3.

Как видно из графика, значение внутрикластерного расстояния уменьшается с увеличением количества нейронов, и, соответственно, количества кластеров, то есть кластеры становятся более компактными, их диаметры уменьшаются. Межкластерное расстояние, напротив, увеличивается при увеличении количества нейронов от 2 до 4, затем при дальнейшем увеличении числа

нейронов межкластерное расстояние уменьшается – более мелкие кластеры располагаются ближе друг к другу. Значение показателя качества кластеризации в данном эксперименте достигает наилучших значений при использовании сети с 4 и 5 нейронами

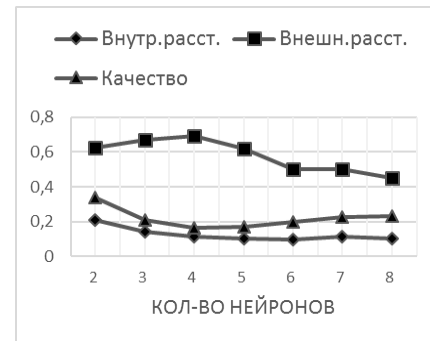


Рис. 3. Изменение показателей качества кластеризации при увеличении количества нейронов

Кластерный анализ является, как правило, начальным этапом изучения данных и извлечения знаний из них. Для дальнейшего анализа требуется произвести содержательное описание полученных кластеров. К примеру, содержательное описание 4 кластеров, полученных в результате использования сети Кохонена с 5 нейронами выглядит следующим образом:

1-й кластер включает клиентов, осуществляющих сделки достаточно редко (с низкой частотой), но проявивших активность сравнительно недавно. По этой причине вероятность осуществления повторных сделок с ними сравнительно высока, их текущая ценность для организации имеет небольшое значение, однако имеет потенциал к повышению.

2-й кластер содержит данные о клиентах, осуществляющих сделки также достаточно редко, к тому же проявивших активность при взаимодействии с организацией достаточно давно. Следовательно, их текущая ценность для организации минимальна.

3-й кластер включает клиентов, осуществляющих сделки с организацией со средней частотой, последний контакт с которыми был произведен сравнительно недавно. В связи с этим они имеют среднюю текущую ценность для организации.

4-й кластер содержит сведения о клиентах, осуществляющих сделки достаточно часто, к тому же проявивших активность сравнительно недавно. Следовательно, клиенты этой группы имеют наивысшую текущую ценность для

организации, им следует уделять наибольшее внимание.

ЗАКЛЮЧЕНИЕ

В данной статье представлены результаты разработки методов сегментации и кластеризации клиентской базы, используемых на начальных этапах процесса управления взаимодействием организации с клиентами. Произведено обоснование актуальности решения задачи разделения клиентской базы на группы и выполнена ее формальная постановка. Предложено использование ряда методов сегментации клиентской базы, произведена их классификация и описание. Использование предложенных методов, в отличие от применения интуитивного эмпирического метода разделения клиентской базы на группы, позволяет выделить сегменты клиентов по заданным критериям в зависимости от задачи, стоящей перед лицом, принимающим решения.

Предложен метод кластеризации клиентской базы, основанный на применении нейронной сети Кохонена, который, в отличие от методов сегментации, позволяет выделять группы клиентов со схожими характеристиками при отсутствии заранее заданных критериев разделения и количестве групп-сегментов.

В рамках дальнейших исследований планируется визуализировать результаты кластеризации с помощью карт Кохонена для ускорения и упрощения процесса принятия решений при управлении взаимодействием с клиентами коммерческих организаций.

СПИСОК ЛИТЕРАТУРЫ

1. **Галямов А. Ф., Тархов С. В.** Поддержка принятия решений при управлении взаимодействием с клиентами на базе центра интегрированного обслуживания // Современные проблемы науки и образования. 2013. № 4. URL: <http://www.science-education.ru/110-9978> (дата обращения: 14.03.2014). [А. Ф. Galyamov, S. V. Tarkhov. Decision support in customer relationship management based on integrated service center [Online], (in Russian), in *Sovremennye Problemy Nauki I obrazovaniya*, vol. 4. Available: <http://www.science-education.ru/110-9978>]

2. **Галямов А. Ф., Тархов С. В.** Модели функционирования и организации хранения информации в системах интегрированного обслуживания клиентов // Вестник УГАТУ. 2012. Т. 16, №3 (48). С. 240–244 [А. Ф. Galyamov, S. V. Tarkhov, "Model of functioning and organization of information storage in integrated customer service systems," (in Russian), *Vestnik UGATU*, vol. 16, no. 3 (48), pp. 240-244, 2011.]

3. **Паклин Н. Б., Орешков В. И.** Бизнес-аналитика: от данных к знаниям. СПб.: Питер, 2013. 704 с. [N. B. Paklin, V. I. Oreshkov. *Business intelligence: from data to knowledge*, (in Russian). St. Petersburg: Piter, 2013.]

ОБ АВТОРАХ

ГАЛЯМОВ Айрат Фаритович, инж.-программист. Дипл. инж. по моделир. и иссл. операций в орг.-техн. системах (УГАТУ, 2008). Канд. техн. наук по упр. в соц. и экон. системах (УГАТУ, 2013). Иссл. в обл. анализа данных в CRM- и ERP-системах.

ТАРХОВ Сергей Владимирович, проф. каф. информатики. Дипл. инж. по технол. машиностроения (УАИ, 1980). Д-р техн. наук по упр. в соц. и экон. системах (УГАТУ, 2010). Иссл. в обл. управления сл. орг.-техн. системами.

METADATA

Title: Customer relationship management of a commercial organization based on methods of segmentation and clustering of customer database.

Authors: A. F. Galyamov¹, S. V. Tarkhov²

Affiliation:

^{1,2} Ufa State Aviation Technical University (USATU).

Email: ¹galyamov.airat@mail.ru, ²tarkhov@inbox.ru

Language: Russian.

Source: Vestnik UGATU (scientific journal of Ufa State Aviation Technical University), vol. 18, no. 4 (65), pp. 149-156, 2014. ISSN 2225-2789 (Online), ISSN 1992-6502 (Print).

Abstract: This paper describes the methods of decision support, used on the initial stages of the customer relationship management of commercial organizations. Considered segmentation and clustering – two classes of methods for separating customer base. Described ABC-analysis, XYZ-analysis, RFM-analysis and NPS-analysis. Kohonen neural network considered as a mechanism for clustering the customer database.

Key words: Customer relationship management; CRM-system; customer database; segmentation; clustering; ABC-analysis; XYZ-analysis; RFM-analysis; NPS-analysis; Kohonen network.

About authors:

GALYAMOV, Airat Faritovich, Cand. of Tech. Sci. (USATU, 2013), engineer of software development (EcoSoft Ltd.), dipl. engineer of modeling and operations research in organizational and technical systems (USATU, 2008). Research in the field of data mining in CRM-and ERP-systems.

TARKHOV, Sergey Vladimirovich, Prof., Dept. of Informatics. Dipl. Engineer of mechanical engineering (Ufa Aviation Institute, 1980). Dr. of Tech. Sci. (UGATU, 2010). Research in the field of organizational management of complex technical systems.