

УДК 007.52

Н.С. Лесная, А.В. Ляховец

## МЕТОДЫ КЛАСТЕРИЗАЦИИ И КЛАССИФИКАЦИИ КЛИЕНТОВ БАНКА И ИХ РЕАЛИЗАЦИЯ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

### Введение

Одной из актуальных задач в сфере автоматизированных банковских систем (АБС) по-прежнему остается задача разработки средств и методов сегментации рынка клиентов банка и анализа кредитоспособности, надежности и перспективности работы с клиентом. Актуальность данных задач связана с бурным развитием банковских структур, стремительным увеличением количества клиентов, расширением спектра услуг, предлагаемых банком. Кредитование сегодня становится массовым явлением. Современная экономическая ситуация подталкивает банки к расширению кредитного предложения. Наряду с понижением кредитной ставки простота оформления и скорость оформления кредита становятся факторами конкурентной борьбы банков за клиентов [1].

В результате сегментации рынка клиентов банка осуществляется возможность определять условия кредитования для каждого конкретного сегмента. С задачей сегментации рынка клиентов банка связана задача классификации новых клиентов и в зависимости от сегмента, к которому относится клиент, предлагать условия кредитования.

Решение данных задач, с одной стороны, даст возможность сформировать условия кредитования для каждого отдельного класса клиентов банка, а с другой — на основании имеющейся осмысленной сегментации появится возможность автоматически предлагать условия кредитования каждому конкретному клиенту в соответствии с его характеристиками.

Следовательно, целью данной работы является нахождение сегментов клиентов банка для оказания услуг кредитования и классификация новых клиентов. Исходя из цели исследования, можно сформулировать следующие задачи исследования: анализ существующих методов сегментации рынка клиентов банка, анализ характеристик клиентов, значимых для определения кредитоспособности клиента, анализ существующих методов кластеризации, анализ существующих методов классификации, возможность применения нейросетевых методов для решения задач сегментации и классификации клиентов банка, построение нейросетевой модели, разработка программного модуля для проведения модельного эксперимента.

### 1. Анализ предметной области

В настоящее время одним из перспективных направлений для развития банка является предо-

ставление услуги кредитования. Для успешного развития банка в данной сфере услуг необходимо учитывать всю совокупность информации о клиентах. Успешное кредитование обусловлено соответствием условий кредита, предложенного клиенту, потребностям и возможностям данного конкретного клиента. Следовательно, для успешной реализации банком основных задач при кредитовании необходимо произвести сегментацию рынка клиентов.

Решение задачи сегментации может быть произведено классическими методами — кластерный анализ, деревья решений, байесовский подход, но из-за динамического изменения характеристик каждого клиента, постоянного добавления важных для банка характеристик и неоднозначности решаемой задачи предлагается использовать методы искусственных нейронных сетей [2].

Для решения первой задачи — сегментирования рынка клиентов банка — предлагается использовать нейронную сеть, решающую задачу кластеризации — самоорганизующиеся карты признаков Кохонена (НС Кохонена). Вторая задача предполагает классификацию новых клиентов на основании их характеристик, решить которую предлагается также с помощью нейросетевого подхода посредством нейронной сети встречного распространения.

Длительный промежуток времени для определения кредитоспособности заемщика и условий кредитования применялись скоринговые карты [3]. Заключение по каждому заемщику основывались на количестве баллов, набранных соискателем кредита. Данный метод является достаточно эффективным, однако в связи с быстроизменяющимися условиями среды кредитования необходим постоянный пересчет скоринговых карт. Операция пересчета является трудоемкой и дорогостоящей, что сильно снижает актуальность использования данного метода.

Для определения наиболее эффективных признаков, позволяющих проводить кластеризацию, проанализированы анкеты, которые каждый новый клиент заполняет при открытии счета.

Из всей совокупности имеющихся факторов были выделены наиболее значимые: возраст; пол; семья; домашний адрес; валюта вклада; сумма депозита; процент; срок; характер выплаты процентов; вид вклада (название); вид рекламы, посредством которой клиент ознакомился с информацией банка.

## 2. Постановка задачи

Задача предоставления выгодных условий кредитования подразделена на две подзадачи.

Первой является задача кластеризации существующих клиентов банка для нахождения сегментов рынка клиентов банка [4]. На основании характеристик клиентов банка, полученных в результате анкетирования, формируется обучающая выборка для НС Кохонена [5].

Нейронная сеть Кохонена, топология которой приведена на рис. 1, позволяет определять не только количество сегментов рынка, а и, самое основное, структуру каждого сегмента рынка клиентов банка, т.е. выявить характерные свойства каждой характеристики в каждом сегменте. При этом сохраняется возможность динамически определять количество сегментов в случае постоянного изменения и обновления количества и характеристик клиентов.

В качестве примера рассмотрим решение задачи сегментации рынка клиентов банка при следующих исходных данных: исследуемая обучающая выборка содержит вектор входных параметров  $X$  длиной 12, по количеству характеристик клиентов, количество элементов 144 (количество клиентов).

В результате решения задачи кластеризации мы должны получить картину о возможных сегментах рынка. Для каждого класса приводятся наиболее характерные значения свойств (для клиентов банка: личные данные, социальное положение, информация об отношениях с банком и т.д.), по которым можно судить о соответствующем сегменте рынка. Полученным сегментам необходимо присвоить номера, а на основании значений свойств присвоить каждому сегменту осмысленные имена.

Классификация новых клиентов банка является второй задачей. Классификация проводится на основании данных, полученных от клиента, но далеко не всегда есть все данные о клиенте. В этом

случае необходимо восстановить неизвестную информацию. При этом в результате решения данной задачи должен быть получен процент принадлежности клиента к каждому из классов. Для решения задачи классификации при наличии неизвестных или отсутствующих значений свойств клиентов, выбрана нейронная сеть встречного распространения.

Для решения указанных задач: кластеризации клиентов с помощью нейронной сети Кохонена и классификации новых клиентов с помощью нейронной сети встречного распространения был разработан соответствующий программный продукт, в котором создана возможность просмотра количества классов, принадлежности конкретного объекта к одному из определенных ранее классов, просмотра значений центроидов классов и средних значений каждого свойства по всей выборке.

## 3. Решение задачи кластеризации

Рассмотрим результаты проведения модельного эксперимента при решении задачи сегментации клиентов банка по предоставлению услуг кредитования. В результате обучения нейронной сети Кохонена производится кластеризация объектов исходной выборки.

Результаты кластеризации исходной выборки, составленной из анкет 144 клиентов банка, представлены в табл. 1. Выявлено, что все 12 характеристик являются значимыми и влияют на принятие решений о выделении сегментов.

На рис. 1 представлена топология нейронной сети Кохонена, которая состоит из одного слоя нейронов и имеет: 12 входных характеристик, которые соответствуют свойствам клиентов, и 4 класса – сегмента рынка клиентов банка [6].

Проанализировав результаты кластеризации, проведенной с помощью нейронной сети Кохонена, выявлено 4 кластера. Каждый кластер является

Таблица 1

Свойства центроидов классов в задаче кластеризации

	Свойства	Средние значения свойств			
		Сегмент 1	Сегмент 2	Сегмент 3	Сегмент 4
1	Пол	муж.	муж.	жен.	жен./муж.
2	Семейное положение	женат/замужем	холост	женат/замужем	холост
3	Дети	есть дети	нет детей	есть дети	нет детей
4	Валюта вклада	UKR	DM	US/UKR	UKR
5	Сумма вклада	1454	3179	1206	1118
6	Процент	25.79%	28.82%	26.02%	20.96%
7	Срок	13 мес.	13 мес.	12 мес.	7 мес.
8	Характер вклада	ежемесячно	в конце срока	в конце срока	ежемесячно
9	Вид вклада	студенческий/ детский	сберегательный	пенсионный/ студенческий	пенсионный/ студенческий
10	Дата открытия счета	15.03.2000	15.03.2000	13.12.1999	03.11.1999
11	Вид рекламы	Радио1/ Радио2	Радио 2/метро	Радио 2/метро	метро
12	Опыт работы с банками	4.94 лет из 7	4.47 лет из 7	5.35 лет из 7	4.59 лет из 7

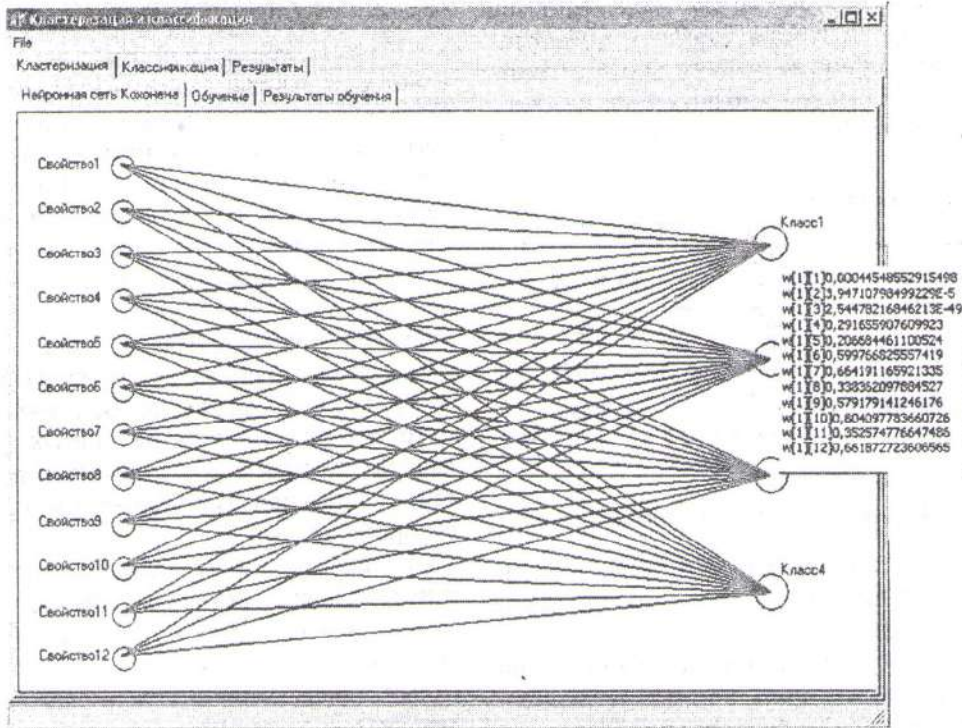


Рис. 1. Топология нейронной сети Кохонена, производящей кластеризацию клиентов

отдельным сегментом рынка банковских услуг кредитования. В соответствии со статистическими характеристиками значений использованных переменных в каждом полученном кластере, предложим следующие названия каждого сегмента:

Сегмент 1 – «клиент ненадежный со средним достатком»;

Сегмент 2 – «клиент ненадежный, обеспеченный»;

Сегмент 3 – «клиент надежный со средним достатком»;

Сегмент 4 – «клиент средней надежности, малообеспеченный».

Вышеперечисленные результаты работы программы после интерпретации и анализа являются

основоположными данными для предоставления определенных услуг и условий кредитования различным группам клиентов банка.

**4. Решение задачи классификации клиентов банка**

Задача классификации решалась с помощью сети встречного распространения. Сеть встречного распространения после ее обучения производит классификацию объектов методом интерполяции, что позволяет восстановить данные, скрытые клиентом, но необходимые банку.

В табл. 2 приведен пример результатов классификации нового клиента с определенными характеристиками (столбец 3).

Таблица 2

**Свойства центроидов классов**

1	Свойства 2	Характеристики нового клиента 3	Средние значения свойств у сегмента 4 4
1	Пол	муж.	жен./муж.
2	Семейное положение	женат/замужем	холост
3	Дети	есть дети	нет детей
4	Валюта вклада	UKR	UKR
5	Сумма вклада	700	1118
6	Процент	22.79%	20.96%
7	Срок	13 мес.	7 мес.
8	Характер вклада	ежемесячно	ежемесячно
9	Вид вклада	пенсионный	пенсионный/студенческий
10	Дата открытия счета	15.08.1999	03.11.1999
11	Вид рекламы	метро	метро
12	Опыт работы с банками	5 лет	4.59 из 7

После подачи его характеристик на вход НС встречного распространения получено распределение вероятностей принадлежности данного клиента ко всем классам (рис. 2), по значению наибольшей вероятности 37% клиент отнесен к сегменту 4.

Класс1	13%	Sex	1,56005707E
Класс2	28%	Fam1	1,636211364
Класс3	22%	Fam2	1,645796054
Класс4	37%	Currency	2,37453191E
		Sum_cf_curr	1251,940917

Рис. 2. Результаты классификации

Средние значения свойств объектов в четвертом сегменте (столбец 4) наиболее точно отражают характеристики данного клиента.

На основании результатов, полученных при классификации, может быть принято решение о возможности предоставления кредита данному клиенту и условий кредитования.

### 5. Программный продукт

Для решения поставленной задачи был разработан программный продукт, который позволяет производить кластеризацию и классификацию клиентов банка, на основании чего банк может строить свои взаимоотношения с клиентами, улучшать уровень обслуживания.

Для решения задачи сегментации клиентов банка в программе создана возможность изменения выборки входных данных, соответственно сегментация может производиться динамически. При сегментации могут быть использованы различные методы и алгоритмы, т.к. для разных выборок применение разных методов и алгоритмов является более результативным. Удобный интерфейс позволяет просмотреть топологию сети, значения центроидов классов, входные значения и результаты.

При решении задачи классификации, данные о новом клиенте могут быть введены как вручную, так и загружены из файла. Реализована возможность работы нейронной сети встречного распространения как в режиме аккредитации, так и в режиме интерполяции [2], что дает возможность просмотра вероятности принадлежности клиента к каждому из классов.

Программа реализована на языке программирования Object Pascal в среде Delphi 7.

### Выводы

В статье предложено решение одной из актуальных задач в сфере автоматизированных банковских систем — задача разработки средств и методов сегментации рынка клиентов банка и анализа кредитоспособности, надежности и перспективности работы с клиентом. Результаты данной работы позволяют предлагать соответствующие условия кредитования каждому конкретному клиенту в соответствии с его характеристиками, возможностями и предпочтениями. Данная задача имеет различные методы решения с использованием различных алгоритмов. Предложенное решение построено на методах искусственных нейронных сетей. Подзадачи были реализованы с помощью нейронных сетей — нейронной сети Кохонена и сети встречного распространения.

Следовательно, в результате данной работы достигнута цель работы: найдены сегменты рынка клиентов банка и произведена классификация новых клиентов. Исходя из цели исследования, были решены следующие задачи исследования: проведен анализ существующих методов сегментации рынка клиентов банка; анализ характеристик клиентов, важных для определения кредитоспособности клиента; анализ существующих методов кластеризации; анализ существующих методов классификации; построена нейросетевая модель; написан программный продукт для проведения эксперимента.

Любая фирма, в том числе и банк, сознает, что ее товары или услуги не могут нравиться сразу всем покупателям. Покупателей слишком много, они отличаются друг от друга своими потребностями и привычками. Каждый банк должен выявить наиболее привлекательные сегменты рынка, которые он в состоянии эффективно обслужить.

**Список литературы:** 1. Зозулев А.В. Сегментирование рынка. М.: Рыбари, 2003. — 178 с. 2. Горбань А., Россиев Д. Нейронные сети на персональном компьютере. Новосибирск: Наука, 1996. — 276 с. 3. Прокурин В.А. // Бизнес и банки. — 2002. — № 30. — С. 11. 4. А.С. Коряк, В.Б. Ренка, Т.Б. Шатовская. Нейросетевое сегментирование рынка банковских услуг. // Труды Международных научно-технических конференций «Интеллектуальные системы» (IEEE ASI'04) и «Интеллектуальные САПР» (CAD-2004): Научное издание в 3-х томах. М.: Изд-во Физико-математической литературы: Т. 1. 2004. — С. 210–213. 5. Kohonen, T. Self-Organizing Maps, Second Edition, Berlin: Springer - Verlag, 1997. 6. Лесная Н.С., Ренка В.Б. Шатовская Т.Б. Интеллектуальный анализ данных. — Харьков: ХНУРЭ, 2003. — 112 с.

Поступила в редколлегию 17.11.2005