

- C. 132-135. URL: file:///C:/Users/User/Downloads/internet-marketing-sushnost-zadachi-preimushestva-i-perspektiv-razvitiya.pdf (дата обращения 25.04.2020).
2. Николаева М.А. Интернет-реклама в продвижении товаров и услуг [Электронный ресурс]: Учебное пособие / М. А. Николаева; Урал.гос. пед. ун-т. – Екатеринбург: [б. и.], 2017. URL: <http://elar.uspu.ru/bitstream/uspu/6428/1/uch00173.pdf> (дата обращения 25.04.2020).
 3. Сатунина М. История контекстной рекламы [Электронный ресурс]. – URL: https://www.e-promo.ru/blog/internet_reklama/istoriya_kontekstnoy_reklamy/ (дата обращения 31.05.2020).
 4. Брагин Л.А., Панкина Т.В. Организация розничной торговли в сети Интернет: учебное пособие для вузов / Л.А. Брагин, Т.В. Панкина. – М.: ФОРУМ: ИНФРА-М, 2014. – С. 120.
 5. Царевский Ф. Яндекс.Директ: как получать прибыль, а не играть в лотерею. 3-е изд., доп. и перераб. — СПб.: Питер, 2017. — 256 с.
 6. ИГХТУ сегодня [Электронный ресурс] – URL: https://www.isuct.ru/sites/default/files/common/isuct_today.pdf (дата обращения 15.05.2020).
 7. Миролюбова А.А., Ксенофонтова О.Л. Научно-исследовательская деятельность вузов: сравнительный анализ и оценка результатов // Известия высших учебных заведений. Серия «Экономика, финансы и управление производством», № 4 (42), 2019, с. 131-138.

References

1. Perova A.A. Internet marketing: essence, tasks, advantages and development prospects [Electronic resource] // Economy and business: theory and practice. - 2018. - No. 9. - S. 132-135. URL: file:///C:/Users/User/Downloads/internet-marketing-sushnost-zadachi-preimushestva-i-perspektiv-razvitiya.pdf (access date 04/25/2020).
2. Nikolaeva M.A. Internet advertising in the promotion of goods and services [Electronic resource]: Textbook / MA Nikolaeva; Ural state ped. un-t. - Yekaterinburg: [b. and.], 2017. URL: <http://elar.uspu.ru/bitstream/uspu/6428/1/uch00173.pdf> (date of access 25.04).
3. Satunina M. History of contextual advertising [Electronic resource]. - URL: https://www.e-promo.ru/blog/internet_reklama/istoriya_kontekstnoy_reklamy/ (date of treatment 05/31/2020)
4. Bragin L.A., Pankina T.V. Organization of retail trade on the Internet: a textbook for universities / L.A. Bragin, T.V. Pankin. - M.: FORUM: INFRA-M, 2014. -- p. 120.
5. Tsarevsky F. Yandex.Direct: how to make a profit, and not play the lottery. 3rd ed., Add. and revised - SPb. : Peter, 2017. -- 256 p.
6. ISUCT today [Electronic resource] - URL: https://www.isuct.ru/sites/default/files/common/isuct_today.pdf (access date 05/15/2020).
7. Миролюбова А.А., Ксенофонтова О.Л. Научно-исследовательская деятельность вузов: сравнительный анализ и оценка результатов // Известия высших учебных заведений. Серия «Экономика, финансы и управление производством», № 4 (42), 2019, с. 131-138.

DOI 10.6060/ivecofin.2020464.509

УДК 004.2

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОЦЕНКИ РИСКОВ ПРИ ВНЕДРЕНИИ НОВОГО КРЕДИТНОГО ПРОДУКТА

Бобков Сергей Петрович

Ивановский государственный химико-технологический университет», пр. Шереметевский, 7, Иваново, 153000, Россия

E-mail: bsp@isuct.ru

Суворов Станислав Вадимович, Орлов Артем Игоревич, Пивнев Егор Алексеевич

Московский политехнический университет, ул. Большая Семеновская, 38, Москва, 107023, Россия

E-mail: ssw1168@mail.ru, artiom65@gmail.com, e.pivnev@yandex.ru

В статье рассмотрены вопросы оценки кредитоспособности частных лиц с использованием кредитного скоринга. Данная система оценки является эффективным подходом к определению уровня риска по конкретному сегменту клиентов. Особенно это касается ситуации запуска кредитной организацией нового кредитного продукта. Основная идея, предложенная в статье, состоит в том, что новые скоринговые карты клиентов создаются на основании уже имеющихся карт путем математической обработки данных. Новизна метода заключается в том, что скоринг делается на основании выделенного подмножества данных клиентов, хранящихся в корпоративном хранилище.

Подход помогает принять решение о выдаче кредита и может быть рекомендован к использованию при кредитовании.

Ключевые слова: нейронные сети, кредитные риски, скоринг, машинное обучение

USING MACHINE LEARNING METHODS TO ASSESS RISKS WHEN IMPLEMENTING A NEW CREDIT PRODUCT

Bobkov Sergey P.

Ivanovo state University of chemistry and technology, Sheremetev avenue, 7, Ivanovo, 153000, Russia

E-mail: bsp@isuct.ru

Suvorov Stanislav V., Orlov Artem I., Pivnev Egor A.

Moscow Polytechnic University, Bolshaya Semenovskaya st., 38, Moscow, 107023, Russia

E-mail: ssw1168@mail.ru, artiom65@gmail.com, e.pivnev@yandex.ru

The article discusses the issues of assessing the creditworthiness of individuals using credit scoring. This rating system is an effective approach to determining the level of risk for a specific customer segment. This is especially true of the situation when a credit institution launches a new credit product. The main idea proposed in the article is that new customer scoring cards are created on the basis of existing cards by mathematical data processing. The novelty of the method lies in the fact that the scoring is done based on a dedicated subset of customer data stored in the corporate storage. The approach helps to make a decision on granting a loan and can be recommended for use in lending.

Keywords: neural networks, credit risks, scoring, machine learning

Введение

На данный момент в России на рынке кредитования существует 870 кредитных организаций и 1969 НКО [9], каждая из которых заинтересована в расширении и продуктовой линейки и регионов присутствия. В данной статье рассмотрены наиболее эффективные подходы к определению уровня риска по конкретному сегменту клиентов, при условии, что данный сегмент является новым для данной компании.

Многие финансовые компании развивают продуктовую линейку и стремятся выйти на новые для себя рынки, однако это невозможно без понимания, кому и с каким результатом будут происходить выдачи [14].

При необходимости расширения бизнеса кредитные организации испытывают трудности с оценкой кредитных рисков на новых для себя клиентах. Это связано с тем, что у каждой кредитной организации есть своя специализация, свой профиль клиента, с которыми они предпочитают работать.

Необходимо понять, как можно качественно определить вероятность дефолта клиента при выходе на новые рынки. Актуальность проблемы очень высокая, так как на рынке кредитования (особенно в среде микрофинансирования) высокая волатильность, и бизнесу необходимо искать новые рынки для продвижения своих услуг.

Теоретическая часть

Известно, что при оценке кредитоспособности частных лиц часто используют систему оценки рисков, называемую кредитным скорингом [4, 5, 6]. Данный подход основан на использовании

численных статистических методов и рекомендуется к использованию при кредитовании на небольшие суммы. При этом математической обработке подвергаются данные, полученные путем заполнения некоторой анкеты. Полученные результаты позволяют принять решение об одобрении или отказе в выдаче кредита.

Положим, есть некая кредитная организация, выдающая долгосрочные залоговые займы, которая выходит на рынок с новым продуктом – PDL (PayDayLoan – заём до зарплаты) [2]. Данный вид кредита подразумевает краткосрочную ссуду до 30 000 руб. на срок до 30 дней (очень актуальный кредитный продукт в настоящее время, что подтверждается недавним увеличением количества микрофинансовых организаций). Для нового продукта прежняя целевая аудитория, как правило, не подходит, и, следовательно, скоринговые карты, созданные ранее, также для нового продукта не подходят.

Процесс должен начинаться проведением маркетинговых исследований для определения социально-демографических характеристик новой целевой аудитории (табл. 1), которая будет основным потребителем нового продукта. Затем необходимо изменить старую скоринговую карту так, чтобы оценивать риск клиента по новому продукту.

Для того, чтобы построить скоринговую карту под новый продукт, необходимо оценить, кто является новой целевой аудиторией данного продукта. Иными словами, необходимо составить портрет клиента, который будет пользоваться новым типом займов. Причём сделать это с точки зрения скоринга –

необходимо найти ключевые отличия в параметрах анкеты клиента, которые участвуют в расчёте балла (балл рассчитывается по т.н. целевой переменной – показателю, определяющему платежеспособность клиента). Для этого понадобится собрать данные по всем клиентам, которые присутствуют в корпоративном хранилище и получить необходимые расчётные поля.

Основная идея подхода, описанного в данной статье, состоит в том, что, используя методы машинного обучения, можно на основании данных одной скоринговой карты создать другую, для другого типа клиента и с другой целевой переменной. В этом и заключается основная новизна предложенного метода. Главное отличие от классического скоринга заключается в том, что из всех данных клиентов, хранящихся в корпоративном хранилище, выделяется некоторое подмножество и на его основании делается скоринг.

Затем необходимо разбить на классы существующий портфель долгосрочных займов и выделить из него тот класс, который будет наиболее похож на новый портрет клиента. Мерой схожести являются евклидовы расстояния между центрами тяжести для каждого класса и эталонного нового клиента в пространстве признаков, образуемых анкетными данными (возраст, пол, социальное положение, доход и т.п.). Тот класс клиентов, который наиболее близок к эталонному клиенту, будет считаться наиболее

подходящим под новый портрет клиента. После получения наиболее подходящего сегмента клиента необходимо получить информацию о том, как сейчас платит этот клиент и обработать его кредитную историю с целью получения новых входных признаков и целевых переменных.

Текущий скоринг настроен на исторических данных для старой целевой аудитории, при этом за целевую переменную брался индикатор наличия просрочки 90+ на 6-й месяц обслуживания клиента. Это оптимальный маркер, т.к. он, во-первых, релевантно показывает платёжеспособность клиентов, а во-вторых, признан стандартами финансовой отчётности.

Очевидно, что для нового продукта такой маркер не подходит, поэтому необходимо *заменить целевую переменную* по тем клиентам, которые будут получены по новым классам. Так как новый продукт подразумевает только один платёж, то имеет смысл брать такой показатель, как просрочка первого (а в нашем случае и единственного) платежа (по графику) на 30 дней.

После того, как будут отобраны платёжные данные (входные признаки и целевая переменная) по новому типу клиента, по ним построить скоринговую карту для нового продукта.

Маркетинговое исследование показало, что средний портрет клиента выглядит так, как показано в таблице 1.

Таблица 1

Показатели старой и новой целевых аудиторий
Table 1. Indicators of old and new target audiences

<i>Показатель</i>	<i>Старая целевая аудитория</i>	<i>Новая целевая аудитория</i>
Средний возраст (лет)	33,4	27,2
Наиболее популярный канал привлечения	Звонок в КЦ по номеру из оффлайн-рекламы	Контекстная реклама в соцсетях
Наиболее популярное семейное положение	Женат/замужем	Холост/не замужем
Наиболее популярное образование	Высшее	Среднее
Наиболее распространённая категория наличия детей	Есть дети	Нет детей
Доля городского населения	63,80%	89,60%

Практическая часть

В данном исследовании использовались, так называемые карты Кохонена [13]. Они были выбраны как один из самых хорошо реализованных на языке R алгоритмов кластеризации. Существенным преимуществом карт Кохонена по сравнению с другими алгоритмами кластеризации является возможность наглядной визуализации кластеров, что также повлияло на выбор

алгоритма кластеризации. Ключевой особенностью карт Кохонена является способность спроецировать многомерные данные на двумерную плоскость. При этом входные векторы с близкими характеристиками на карте будут располагаться рядом. Таким образом, входные векторы со схожими характеристиками будут образовывать кластеры. Количество кластеров определяется автоматически в процессе обучения, что тоже

выгодно отличает карты Кохонена от других методов кластеризации.

Ячейки, визуализирующие входные признаки, в картах Кохонена обычно раскрашиваются разными цветами или могут иметь градации серого цвета. Цвет каждой ячейки считается по каждому входному признаку отдельно. Весь диапазон изменения входного признака сопоставляется с диапазоном изменения цвета. Например, признак возраста человека может меняться в диапазоне от 1 до 120 лет. Градации же серого меняются в диапазоне от 0 до 255. Таким образом, значение 120 лет можно соотнести с диапазоном цвета [253, 255], 119 лет с диапазоном [250, 252] и т.д. Ту же самую процедуру повторяем с другим входным признаком, например, с доходом. Получаем для каждой ячейки уже другой цвет или градацию серого. После чего для каждой ячейки находим среднее значение цвета.

Все вычисления производились на языке R версии 3.5.0.

Результаты кластеризации представлены на рисунке 1.

На данном рисунке видно, что необходимый сегмент клиентов выделяется в обособленную группу (справа внизу) и которую необходимо использовать для построения новой скоринговой карты.

Кластеры клиентов

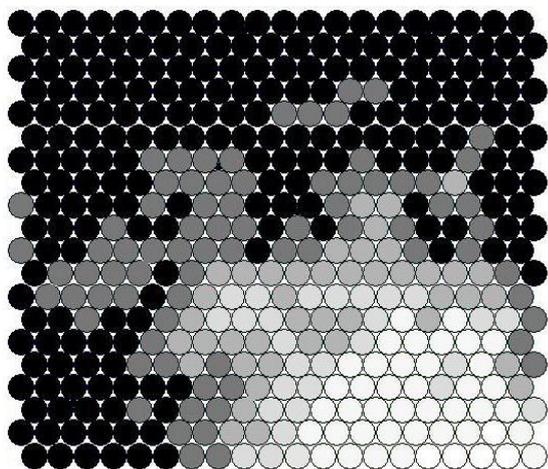


Рисунок 1. Результаты кластеризации с использованием карт Кохонена (кластеры клиентов)

Figure 1. Results of clustering using Kohonen maps (customer clusters)

Полученные данные новой целевой аудитории, а также переопределённая целевая переменная (с риска на 6-й месяц обслуживания на риск на 2-й месяц обслуживания) позволяют построить необходимую скоринговую модель под новый

продукт.

Характеристики сравниваемых скоринговых моделей представлены в таблице 2.

Коэффициент Gini и статистика Колмогорова-Смирнова – это классические метрики качества скоринговой модели.

Коэффициент Gini позволяет судить о дискриминирующей силе модели, тогда как статистика Колмогорова-Смирнова – максимальное «расстояние» между функциями распределения по баллу хороших и плохих. Из таблицы 2 видно, что предлагаемая скоринговая модель по качеству не значительно уступает модели для первого продукта, а значит, также имеет высокую разделяющую силу.

Результаты окончательного сравнения итоговых финансовых показателей анализируемых моделей приведены в таблице 3.

Таблица 2

Итоговые характеристики моделей
Table 2. The final characteristics of the models

Показатель	Скоринговая модель для первого продукта	Скоринговая модель для нового продукта
Количество наблюдений	17276	5784
Коэффициент Gini	0,51	0,48
Статистика Колмогорова-Смирнова	49,2	42,1

Таблица 3

Итоговые финансовые показатели моделей
Table 3. The final financial characteristics of the models

Показатель	Скоринговая модель для первого продукта	Скоринговая модель для нового продукта
Прогнозируемые конечные потери (% от суммы выдачи)	16,2%	23,6%
Доля клиентов, просрочивших первый платёж на 30 дней	6,7%	12,3%

Можно отметить, что приведенные показатели позволяют оценить финансовый итог выхода компании на новый рынок, а также построить финансовую модель для бюджетирования и прогнозирования необходимых финансовых показателей.

Заключение

В данной работе был предложен подход, позволяющий, применяя методы машинного обучения, использовать данные существующей скоринговой карты для создания другой карты.

При этом созданная карта будет предназначена для нового типа клиента (продукта) и будет отличаться целевой переменной. Цель достигается путем анализа данных клиентов, хранящихся в корпоративном хранилище и выделения из них класса, наиболее подходящего под новый портрет клиента. В этом случае также определяется индикатор (маркер) платежеспособности клиента. В качестве инструмента были использованы нейронные сети в форме карт Кохонена. При тестировании предложенный метод показал высокую эффективность.

Литература

1. Кредитное бюро Эквифакс [Электронный ресурс]. - URL: <https://www.equifax.ru> (дата обращения: 12.04.2019).
2. Мировая экономика финансы и инвестиции PDL [Электронный ресурс]. - URL: <http://www.globfin.ru/articles/banks/chto-takoe-pdl-zajmy-i-gde-ih-vzyat.htm> (дата обращения: 14.04.2019).
3. Новичкова О.В. Управление кредитным риском организации // Инновационная наука. - 2015. - № 11-1. - С. 134-137.
4. Пищулин А. Кредитный скоринг. Не всё так страшно. [Электронный ресурс]. - URL: <http://factoringpro.ru/index.php/credit-scoring-statya/408-skoring-statya-kredit> (дата обращения: 14.03.2019).
5. Севрук В.Т. Банковские риски - М.: Дело, 1995. - 72 с.
6. Спичёва А.Ф. К вопросу об управлении кредитным риском // Наука. Образование. Личность. - 2013. - № 1. - С. 144-148.
7. ФССП России [Электронный ресурс]. - URL: <http://fssprus.ru/statistics/> (дата обращения: 12.04.2019).
8. Хандруев А.А. Управление рисками банков: Научно-практический аспект // Деньги и кредит. - 1997. - № 6. - С. 12.
9. ЦБ РФ [Электронный ресурс]. - URL: <http://cbr.ru> (дата обращения: 14.04.2019).
10. Шамхалова Э.А. Некоторые особенности управления кредитным риском // Горизонты экономики. - 2014. - № 6-2 (19). - С. 93-97.
11. Шунина Ю.С., Алексеева В.А., Клячкин В.Н. Прогнозирование кредитоспособности клиентов на основе методов машинного обучения // Финансы и кредит. - 2015. - № 27 (651). - С. 2-12.
12. Philip Kotler: Marketing Essentials, 1984. / Филип Котлер: Основы маркетинга. Перевод на русский язык: В. Б. Бобров. — М., 1990.
13. Introduction to Self-Organizing Maps in R [Электронный ресурс]. - URL: https://clarkdatalabs.github.io/soms/SOM_NBA (Дата обращения: 15.09.2019).
14. Бибикова Е.А., Валинурова А.А., Третьякова Ю.С. Оценка эффективности внедрения инновационного банковского продукта // Известия высших учебных заведений. Серия «Экономика, финансы и управление производством», № 1 (39), 2019, с. 3-10.

References

1. Credit Bureau Equifax [Electronic resource]. - URL: <https://www.equifax.ru> (date of access: 12.04.2019).
2. World economy finance and investments PDL [Electronic resource]. - URL: <http://www.globfin.ru/articles/banks/chto-takoe-pdl-zajmy-i-gde-ih-vzyat.htm> (date of access: 14.04.2019).
3. Novichkova O.V. Organization's credit risk management // Innovative Science. - 2015. - No. 11-1. - S. 134-137.
4. Pishchulin A. Credit scoring. It's not all that bad. [Electronic resource]. - URL: <http://factoringpro.ru/index.php/credit-scoring-statya/408-skoring-statya-kredit> (date of access: 03/14/2019).
5. Sevruk V.T. Banking risks - M.: Delo, 1995. -- 72 p.
6. Spicheva A.F. On the issue of credit risk management // Science. Education. Personality. - 2013. - No. 1. - S. 144-148.
7. FSSP of Russia [Electronic resource]. - URL: <http://fssprus.ru/statistics/> (date of access: 12.04.2019).
8. Khandruev A.A. Banking Risk Management: Scientific and Practical Aspect // Money and Credit. - 1997. - No. 6. - P. 12.
9. Central Bank of the Russian Federation [Electronic resource]. - URL: <http://cbr.ru> (date of access: 14.04.2019).
10. Shamkhalova E.A. Some features of credit risk management // Horizons of Economics. - 2014. - No. 6-2 (19). - S. 93-97.
11. Shunina Yu.S., Alekseeva VA, Klyachkin VN Predicting the creditworthiness of clients based on machine learning methods // Finance and Credit. - 2015. - No. 27 (651). - S. 2-12.
12. Philip Kotler: Marketing Essentials, 1984. / Philip Kotler: Fundamentals of Marketing. Translation into Russian: V. B. Bobrov. - M., 1990.
13. Introduction to Self-Organizing Maps in R [Electronic resource]. - URL: https://clarkdatalabs.github.io/soms/SOM_NBA (Date of access: 15.09.2019).
14. Бибикова Е.А., Валинурова А.А., Третьякова Ю.С. Оценка эффективности внедрения инновационного банковского продукта // Известия высших учебных заведений. Серия «Экономика, финансы и управление производством», № 1 (39), 2019, с. 3-10.