

На правах рукописи



Романюк Кирилл Андреевич

**Разработка метода и модели классификации
объектов по непрерывной шкале
для принятия решений о параметрах
кредитования физических лиц**

Специальность 08.00.13 —
«Математические и инструментальные методы экономики»

Автореферат
диссертации на соискание учёной степени
кандидата экономических наук

Санкт-Петербург — 2016

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Санкт-Петербургский политехнический университет Петра Великого».

Научный руководитель: доктор экономических наук, профессор
Ильин Игорь Васильевич

Официальные оппоненты: **Воронцовский Алексей Владимирович**
доктор экономических наук, профессор,
ФГБОУ ВО «Санкт-Петербургский государственный университет», профессор кафедры экономической кибернетики
Лужбин Алексей Анатольевич
кандидат экономических наук, ПАО «Выборг-банк», руководитель группы малого и среднего бизнеса службы кредитного андеррайтинга

Ведущая организация: ФГБОУ ВО «Санкт-Петербургский государственный экономический университет»

Защита состоится «29» сентября 2016 года в 14:00 часов на заседании диссертационного совета Д 999.056.02 при ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого» и при ФГАОУ ВО «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики», по адресу: 195251, Санкт-Петербург, ул. Политехническая, д. 29, III учебный корпус, ауд. 506.

С диссертацией можно ознакомиться в Фундаментальной библиотеке и на сайте ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого» <http://www.spbstu.ru/science/defences.html>, в библиотеке и на сайте ФГАОУ ВО «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики» <http://fppo.ifmo.ru>.

Автореферат разослан «___» июня 2016 г.

Ученый секретарь диссертационного совета,
доктор экономических наук,
доцент



Калинина Ольга Владимировна

Общая характеристика работы

Актуальность темы диссертационной работы. За последние два года, после введения рядом стран ограничительных мер в отношении российской экономики, и снижении финансовых поступлений в страну, после падения цен на нефть, число отозванных лицензий у банков в России многократно увеличилось. Данный факт свидетельствует о низком качестве управления кредитным риском, то есть банковская система не способна своевременно адаптироваться относительно колебаний рыночной конъюнктуры. Рынок потребительского кредитования является весомой частью банковской деятельности. На рынке потребительского кредитования существует проблема усреднённого ценообразования. Физические лица, которые подают заявку на один тип кредита, получают одинаковую процентную ставку вне зависимости от кредитоспособности. Причиной является ограниченность методов классификации объектов, используемых при определении кредитоспособности физических лиц. Большинство подобных методов позволяет классифицировать физических лиц по бинарной шкале: «кредитоспособен», «некредитоспособен». «Некредитоспособные» клиенты банка не получают кредит, как обладатели наибольшего кредитного риска. «Кредитоспособные» клиенты банка получают кредит по единой процентной ставке. Указанная ограниченность методов классификации является препятствием для автоматизации взаимодействия банка с клиентом, которая необходима, как для снижения затрат на ведение банковской деятельности, так и для своевременного управления кредитным риском, а так же порождает неблагоприятный отбор клиентов.

Множество значений, которое может принимать кредитоспособность, является континуумом. Для дифференциации процентной ставки физическому лицу в соответствии с действительной кредитоспособностью классификация физических лиц по кредитоспособности должна осуществляться по непрерывной шкале. В связи с этим актуальна разработка систем поддержки принятия решений для дифференциации процентной ставки по кредиту физическим лицам, и в особенности моделей и методов классификации физических лиц по кредитоспособности по непрерывной шкале, обеспечивающих функционирование данных систем. Дифференцирование условий кредитования в зависимости от кредитоспособности клиента позволит банкам качественнее распределять финансовые ресурсы, снижая потери по ссудам и риск отзыва лицензии при колебаниях рыночной конъюнктуры.

Степень разработанности проблемы. Стандартная модель классификации объектов при анализе кредитоспособности физических лиц состоит в отображении частной информации клиента на бинарную шкалу: «кредитоспособен», «некредитоспособен». Модели классификации объектов для анализа кредитоспособности физических лиц с числом классов больше, чем два предложены в работах Б. Чи, К. Шу, М. Хо И. Хатзилыгероудис, Д.

Прентзас, К. Ксинуй, К. Жонг и П. Маррикату, К. Шанипугари. Увеличение числа классов при фиксированных границах в пределе даёт непрерывную шкалу для классификации. В текущей работе представлена модель классификации объектов по непрерывной шкале для анализа кредитоспособности физических лиц.

Вклад в развитие математических методов, обеспечивающих более широкие возможности в классификации физических лиц по кредитоспособности и в решение сопутствующих вопросов, внесли Т. Беллотти, Д. Крук, Т. Харрис, А. Кашман, С. Ю. Чанг, А. Капоторти, Г. Ю. Силкина, В. Хенлей, Ч. Л. Хуанг, А. Лаха, Р. Малхотра, Б. В. Чи, В. Гарсия, А. И. Маркес, Д. С. Санчес, П. Даненас, Ю. С. Ким, В. В. Глухов, С. Наеем, С. Орески, А. Г. Перевозчиков, В. С. Десай, Р. Милерис, С. Ю. Сон, Д. Вест, Р. Сетионо.

Существенный вклад в развитие теоретических и методологических основ классификации физических лиц по кредитоспособности внесли Дэвид Хэнд, Барт Байесенс и Лин Томас. В работах перечисленных специалистов можно увидеть, что стандартная модель классификации физических лиц по кредитоспособности исчерпала себя. Д. Хэнд посвятил статью тому, что новые, более сложные методы классификации зачастую создают иллюзию прогресса, не увеличивая точности.

Вклад в развитие метода рандомизированных сводных показателей, используемого в работе, внесли отечественные и зарубежные специалисты Н. В. Хованов, Д. Н. Колесов, М. С. Юдаева, С. Е. Шмелёв, Б. Вуциевич, Н. Афган, К. Конту, С. Ринне.

Кроме того, усреднение процентной ставки по кредиту физическим лицам порождает негативные эффекты асимметрии информации, в частности, неблагоприятный отбор клиентов на рынке кредитования физических лиц. Вклад в развитие основ, связанных с асимметрией информации, внесли С. Д. Гроссман, М. Ротшильд, Ю. Фама, Р. Бабб, Т. Джаппелли, М. Пагано, А. В. Воронцовский, Дж. Акерлоф, Дж. Стиглиц, Д. Карлан и Д. Цинман оценили практический эффект асимметрии информации на рынке потребительского кредитования. В текущей работе рассматривается вопрос, как классификацию физических лиц по кредитоспособности по непрерывной шкале можно использовать для снижения негативных эффектов асимметрии информации на рынке кредитования физических лиц.

Целью работы является разработка модели и метода классификации объектов по непрерывной шкале при кредитовании физических лиц.

Для достижения данной цели необходимо решить следующие **задачи**:

1. Рассмотреть воздействие применяемых моделей классификации объектов при выдаче кредита физическим лицам на рынок потребительского кредитования, в частности, на проблему асимметрии информации для выявления подходящей модели классификации последних по кредитоспособности.

2. Проанализировать существующие модели и методы классификации объектов, чтобы сформулировать условия для нахождения оптимального результата классификации.
3. Обосновать предельный переход от моделей классификации объектов по дискретной шкале к модели по непрерывной шкале.
4. Апробировать разработанный метод по базе кредитных историй для классификации физических лиц по кредитоспособности.
5. Разработать систему поддержки принятия решений для вычисления стоимостной оценки результата классификации физических лиц по кредитоспособности.
6. Оценить экономический эффект от внедрения разработанного метода и системы поддержки принятия решений.
7. Проанализировать для каких банков рентабельно внедрение разрабатываемой системы поддержки принятия решений.
8. Определить порядок действий для автоматизированной корректировки атрибутов контракта между банком и физическим лицом на основании классификации последнего по кредитоспособности.

Объектом исследования являются коммерческие банки, принимающие управленческие решения о выдаче кредитов физическим лицам. **Предметом** исследования является математический аппарат классификации физических лиц – потенциальных заемщиков банка как основы дифференциации условий кредита.

Теоретической основой исследования является теория рынков с асимметричным распределением информации, а так же результаты, изложенные в трудах ведущих мировых специалистов по методам классификации объектов, в том числе при кредитовании физических лиц.

Методической основой исследования является метод рандомизированных сводных показателей, методы теории нечётких множеств и метод Монте-Карло.

Информационную базу исследования составляют данные ЦБ РФ, а так же база кредитных историй, доступная по электронному адресу <https://archive.ics.uci.edu/ml/datasets.html>.

Научная новизна исследования состоит в разработке модели и метода классификации объектов по непрерывной шкале, которые применены для анализа кредитоспособности физических лиц, а так же в разработке системы поддержки принятия решений для рационализации банковского управления процентной ставкой по кредиту. Более подробно научная новизна раскрывается в **основных положениях, выносимых на защиту:**

1. Предложена модель классификации объектов по непрерывной шкале при кредитовании физических лиц.
2. Сформулирована задача минимизации доли потерь по ссудам банка на основе управления показателем точности классификации при ограничениях на значения весовых коэффициентов.
3. Разработан метод классификации объектов с использованием метода рандомизированных сводных показателей и методов теории нечётких множеств, и обоснована возможность его применения при анализе кредитоспособности физических лиц.
4. Разработана система поддержки принятия решений по установлению процентной ставки на основе классификации физических лиц по кредитоспособности по непрерывной шкале.
5. Предложен алгоритм параметрической оценки атрибутов кредита при многопериодном взаимодействии с клиентом на основе классификации физического лица по кредитоспособности.

Теоретическая значимость состоит в развитии математического аппарата классификации объектов, а так же в выявлении варианта решения одной из проблем асимметрии информации, а именно неблагоприятного отбора клиентов на рынке кредитования физических лиц.

Практическая значимость заключается в том, что применение результатов исследования обеспечит банку следующее: снижение потерь по кредитам на единицу выданных займов; возможность визуализировать допустимые комбинации условий кредита в соответствии с классом кредитоспособности физического лица и корректировать данные условия по динамике вычисленного класса; снижение отклонения процентной ставки по кредиту от потерь по кредиту, увеличивая финансовую стабильность к колебаниям рыночной конъюнктуры.

Достоверность полученных результатов обеспечивается использованием актуальных статистических данных и существенного количества международных рецензируемых научных изданий, а так же логической стройностью разработанных модели и метода.

Область исследований по паспорту специальности 08.00.13, математические и инструментальные методы экономики, соответствует пунктам:

- 1.6. Математический анализ и моделирование процессов в финансовом секторе экономики, развитие метода финансовой математики и актуарных расчетов.
- 2.3. Разработка систем поддержки принятия решений для рационализации организационных структур и оптимизации управления экономикой на всех уровнях.

Апробация работы. Основные результаты работы докладывались на международных конференциях, проводимых на английском языке:

1. EAIS, Франция, Дуэ, НИИ Минес, 1-3 декабря 2015 г.;
2. IntelliSys, Великобритания, Лондон, 10-11 ноября 2015 г.;
3. RuSSIR YSC, Санкт-Петербург, НИИ ВШЭ, 25-26 августа 2015 г.;
4. SPVCEMF, Санкт-Петербург, СПбПУ, 18-20 марта 2015 г.

Результаты работы обсуждались при личной встрече с членом британской академии наук, обладателем Ордена Британской Империи за вклад в развитие науки, профессором Дэвидом Хэндом, который является главным научным руководителем в финансовой организации Винтон Капитал (Winton Capital). Профессор Хэнд положительно охарактеризовал результаты работы и дал рекомендацию для начальника команды по оценке кредитного риска в Имперском Колледже Лондона, доктору Тони Беллоти, с которыми результаты работы так же обсуждались при личной встрече.

Практическая апробация работы осуществлена в ПАО «Банк «Санкт-Петербург».

Публикации. Основные результаты по теме диссертации опубликованы в 7-ми рецензируемых научных изданиях, которые рекомендованы Министерством образования и науки Российской Федерации (в которых должны быть опубликованы основные научные результаты диссертации на соискание ученой степени), в том числе 3, которые входят в реферативную базу данных Scopus.

Объем и структура работы. Диссертация состоит из введения, трёх глав, заключения, списка литературы и приложения. Объем диссертации составляет 184 страницы с 26 рисунками и 29 таблицами. Список литературы содержит 150 наименований.

Описание основных положений работы

1. Предложена модель классификации объектов по непрерывной шкале при кредитовании физических лиц.

При принятии решения о выдаче кредита физическим лицам банк классифицирует последних по кредитоспособности. Проблема состоит в том, что применение существующих моделей классификации физических лиц по кредитоспособности ограничивает возможности банка в дифференциации процентной ставки по кредиту и, следовательно, препятствует снижению риска банкротства последнего при колебаниях рыночной конъюнктуры за счёт подобной дифференциации. Для решения данной проблемы предложена модель классификации объектов.

Модель классификации объектов по непрерывной шкале состоит в следующем. Каждый из исследуемых объектов описывается вектором характеристик $x = \{x_1, x_2, \dots, x_m\}$ из множества допустимых значений X . На основании векторов характеристик объекты классифицируются по непрерывной шкале, для чего строится показатель $Q(x, w) \in [0, 1]$, который отображает класс объекта, где $w = \{w_1, w_2, \dots, w_m\}$ – вектор параметров из множества допустимых значений W .

Построенный показатель $Q(x, w)$ банк может использовать при кредитовании физических лиц, составив векторы характеристик объектов из частной информации физических лиц (возраст, доход и т.д.). С помощью данного показателя банк может решить следующие задачи:

1. определить группу клиентов, которым следует отказать в выдаче кредита;
2. дифференцировать процентную ставку по кредиту остальным клиентам в соответствии с кредитоспособностью.

В группу клиентов, которые получают отказ в выдаче кредита, попадают те физические лица, чья кредитоспособность ниже минимально допустимого значения Q_{min} : $\{j : Q_j < Q_{min}\}$, где $Q_j = Q(x^{(j)}, w)$ – кредитоспособность физического лица с набором характеристик $x^{(j)} = \{x_1^{(j)}, x_2^{(j)}, \dots, x_m^{(j)}\}$.

Процентную ставку по кредиту можно представить, как сумму процентной ставки без кредитного риска P_{min} и премии за кредитный риск c : $P = P_{min} + c$. Способ выражения премии за кредитный риск через класс кредитоспособности представлен в описании положения 4 основных научных результатов работы: $c(Q) = \sum_{i=0}^k a_i Q^i$, $a_i \in R, \forall i = 0, 1, 2, \dots, k$. Для подобного представления искомый показатель кредитоспособности $Q(x, w)$ необходимо строить на основе анализа потерь по ссудам, выданных физическим лицам. Таким образом, процентную ставку по кредиту для j -го клиента можно определить через класс кредитоспособности: $P(Q_j) = P_{min} + \sum_{i=0}^k a_i Q_j^i$.

Премия за кредитный риск может принимать значения в диапазоне: $c(Q) \in [0, c_{max}]$, где $c_{max} = c(Q_{min})$. Следовательно, процентная ставка по кредиту так же может принимать значения из некоторого заданного диапазона (по непрерывной шкале): $P(Q) \in [P_{min}, P_{min} + c_{max}]$, в отличие от существующих моделей.

Наиболее распространённая модель классификации физических лиц по кредитоспособности является бинарной: $Q^{(2)}(x, w) \in \{0, 1\}$. Результат подобной классификации банк использует, чтобы принять решение выдавать кредит клиенту по единой процентной ставке или отказать в выдаче кредита. Если банк устанавливает единую процентную ставку, то порождается неблагоприятный отбор клиентов. Другими словами, клиенты с высокой кредитоспособностью заинтересованы брать кредит в банке, в котором процентная ставка дифференцирована в соответствии с кредитоспособностью, что

обеспечивает более высокую конкурентоспособность банка при использовании предложенной модели.

Существуют модели, в которых классификация физических лиц по кредитоспособности осуществляется с использованием трех и пяти классов, соответственно: $Q^{(3)}(x, w) \in \{0, \frac{1}{2}, 1\}$, $Q^{(5)}(x, w) \in \{0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, 1\}$. При использовании указанных моделей банк может дифференцировать процентную ставку на две ($P \in \{P_1, P_2\}$) и четыре категории ($P \in \{P_1, P_2, P_3, P_4\}$), соответственно. В общем виде модель классификации физических лиц по кредитоспособности может быть задана с помощью функции, отображающей частную информацию на числовую шкалу с n градациями:

$$Q^{(n)}(x, w) \in \{0, \frac{1}{n-1}, \frac{2}{n-1}, \dots, \frac{n-2}{n-1}, 1\}.$$

Предложенную модель можно получить, как предельный переход от существующих моделей классификации физических лиц по кредитоспособности при увеличении числа классов до плюс бесконечности, что сформулировано в виде следующей теоремы.

Теорема. Для последовательности функций $\{Q^{(n)}(x, w)\}$ при фиксированном $w \in W$ функция $Q(x, w)$ является пределом при увеличении числа n до плюс бесконечности при любом наборе исходных характеристик $x \in X$.

Доказательство. Пусть x – набор характеристик из множества X . Тогда разность оценки кредитоспособности по дискретной и непрерывной функции меньше величины градации используемой шкалы:

$$|Q^{(n)}(x) - Q(x)| < \frac{1}{n-1}, \forall x \in X.$$

При увеличении числа градаций величина градации стремится к нулю:

$$\lim_{n \rightarrow +\infty} \frac{1}{n-1} = 0.$$

Таким образом, при любом $\varepsilon > 0$ существует число N_ε такое, что при любом $n > N_\varepsilon$ выполняется неравенство $|Q^{(n)}(x) - Q(x)| < \varepsilon$. Следовательно, последовательность функций равномерно сходится к функции Q :

$$\{Q^{(n)}\} \Rightarrow Q \iff \lim_{n \rightarrow +\infty} \sup_{x \in X} |Q^{(n)}(x) - Q(x)| = 0.$$

Каждый объект по данной модели получает индивидуальный класс. В случае применения данной модели при классификации физических лиц по кредитоспособности каждое физическое лицо получит индивидуальный класс кредитоспособности. Любая классификация по кредитоспособности по дискретной шкале может быть получена посредством усреднения информации из предложенной модели. Получается, что после осуществления классификации физических лиц по кредитоспособности по непрерывной шкале банк может задать любое количество классов клиентов, для которых будет установлено различное значение процентных ставок по кредиту, что обеспечива-

ет более высокую финансовую стабильность банка при колебаниях рыночной конъюнктуры.

2. Сформулирована задача минимизации доли потерь по ссудам банка на основе управления показателем точности классификации при ограничениях на значения весовых коэффициентов.

При отсеве некредитоспособных клиентов существенное значение имеет степень соответствия результата классификации последних по кредитоспособности заданному критерию, что оценивается показателем общей точности классификатора (Total Accuracy). Более высокое значение данного показателя обеспечивает более низкую величину потерь по ссудам на единицу выданных средств: $c(TA) = 1 - TA$. В связи с этим целесообразно настроить параметры предложенной модели таким образом, чтобы доля потерь по ссудам принимала наименьшее значение при некоторых ограничениях:

$$c(TA(Q(w))) \rightarrow \min,$$

$$\left\{ \begin{array}{l} w_i - g_i \leq 0, i = 1, 2, \dots, m, \\ w_i - f_i \geq 0, i = 1, 2, \dots, m, \\ w_i - w_j > 0, \text{ при } L(i, j) = 1, \\ w_i - w_j = 0, \text{ при } E(i, j) = 1, \\ 0 \leq f_i \leq g_i \leq 1, i = 1, 2, \dots, m, \\ \sum_{i=1}^m w_i = 1, w_j \geq 0, j = 1, 2, \dots, m \end{array} \right.$$

где $c(TA(Q(w))) = 1 - TA(Q(w))$, $w = (w_1, w_2, \dots, w_m)$ – вектор весовых коэффициентов (параметров), $TA(Q(w)) = TPR(Q(w)) + TNR(Q(w))$; $TA(Q(w))$ – показатель общей точности классификации при векторе w ; $TPR(Q(w))$ – доля верно классифицированных положительных наблюдений при векторе w ; $TNR(Q(w))$ – доля верно классифицированных отрицательных наблюдений при векторе w ; $L_{m \times m}$, $E_{m \times m}$ – матрицы, состоящие из нулей и единиц, если $L(i, j) = 1$, то $w_i > w_j$, если $E(i, j) = 1$, то $w_i = w_j$, если $L(i, j) = 0$ и $E(i, j) = 0$, то дополнительное условие для пары w_i и w_j отсутствует.

В терминах метода рандомизированных сводных показателей, приведенная оптимизационная задача может быть представлена в следующем виде: $c(TA(Q(w))) \rightarrow \min$, при ограничениях $w \in OI \cap UI \cap NC$, где

- OI – нечисловая информация, задающая порядок доминирования весовых коэффициентов, $OI = \{w_l > w_k, w_e = w_s, l, k, e, s \in \{1, 2, \dots, m\}\}$;
- UI – неточная информация, задающая числовой диапазон варьирования весовых коэффициентов, $UI = \{0 \leq f_i \leq w_i \leq g_i \leq 1, i = 1, 2, \dots, m\}$;
- NC – множество коэффициентов, сформированное нормирующим условием, $NC = \{w = (w_1, w_2, \dots, w_m) : \sum_{i=1}^m w_i = 1, w_j \geq 0, j = 1, 2, \dots, m\}$.

В данной задаче максимизируется значение показателя общей точности классификатора (ТА) при ограничениях представленных дополнительной информацией о значимости характеристик и нормирующим условием. Так же задачу можно сформулировать, как минимизацию ошибки классификации при тех же ограничениях. Стоит отметить, что целевая функция дискретна, что усложняет поиск решения. Решение данной задачи осуществлено с использованием инструментария метода рандомизированных сводных показателей и методов теории нечетких множеств, что описано в следующем положении основных научных результатов работы.

3. Разработан метод классификации объектов с использованием метода рандомизированных сводных показателей и методов теории нечётких множеств, и обоснована возможность его применения при анализе кредитоспособности физических лиц.

Основные этапы метода состоят в следующем:

1. Составить вектор характеристик объекта, физического лица:

$$x = (x_1, x_2, \dots, x_m).$$
2. Оценить характеристики в диапазоне от 0 до 1: $q_i(x_i) \in [0, 1], i = 1, 2, \dots, m.$
3. Определить весовые коэффициенты, которые задают значимость каждой характеристике: $w = (w_1, w_2, \dots, w_m), w_i \geq 0, \sum_{i=1}^m w_i = 1.$
4. Вычислить значение агрегирующей функции.

В явном виде класс кредитоспособности для i -го выглядит следующим образом: $Q_j = Q(q^{(j)}; w) = Q(q_1(x_1^{(j)}), q_2(x_2^{(j)}), \dots, q_m(x_m^{(j)}); w_1, w_2, \dots, w_m) \in [0, 1].$ Оценка исходных характеристик осуществляется с использованием теории нечётких множеств. В качестве агрегирующей функции в методе рандомизированных сводных показателей часто используется функция средневзвешенного арифметического. Значение весовых коэффициентов определяется путём рандомизации на множестве допустимых значений:

$$W(m, n) = \{w^{(t)} = (w_1^{(t)}, w_2^{(t)}, \dots, w_m^{(t)}), w_i^{(t)} \in \{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}, \\ w_1^{(t)} + w_2^{(t)} + \dots + w_m^{(t)} = 1, i = 1, 2, \dots, m, t = 1, 2, \dots, N(m, n)\},$$

где $W(m, n)$ – множество допустимых комбинаций весовых коэффициентов, $N(m, n)$ – число допустимых комбинаций весовых коэффициентов.

Область допустимых значений определяется с использованием дополнительной, нечисловой и неточной информации, которая описана в предыдущем положении основных научных результатов работы. Данная информация может быть получена как статистическими, так и экспертными методами.

В совокупности условия по дополнительной информации образуют область допустимых значений для весовых коэффициентов, графически представляемую в форме многогранника. Весовые коэффициенты рандомизируются с некоторым распределением, в частности, равномерным по данной области. В результате рандомизации весовые коэффициенты и оценка объекта становятся случайными величинами: $\tilde{w} = (\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_m)$, $\tilde{Q} = \tilde{Q}(q(x); \tilde{w})$.

Анализируемая выборка кредитных историй разбивается на обучающее и тестовое множество наблюдений. По обучающему множеству определяются значения переменных, которые максимизируют значение целевой функции. Затем данные значения используются для классификации физических лиц по кредитоспособности на тестовом множестве.

На этапе оценки исходных характеристик использованы методы теории нечётких множеств, а именно: использована импликация Мамдами с максиминным методом получения результирующей функции принадлежности, где дефазификация осуществлена с помощью метода центра тяжести.

При сравнении методов классификации, апробированных на используемой базе кредитных историй по вероятности верно классифицированных наблюдений, составлена Таблица 1.

Таблица 1: Точность методов классификации

Название	Точность (Total accuracy)
метод К ближайших соседей	62,0%
метод опорных векторов	63,5%
разработанный метод	65,2%

В сравнении с методом опорных векторов и методом К ближайших соседей разработанный метод показал более высокую точность по вероятности верно предсказанных исходов. При внедрении разработанного метода в банк, не меняя ценообразования кредита, положительный эффект ожидается от снижения потерь по ссудам за счёт более точной классификации потенциальных клиентов по кредитоспособности. Для банка с портфелем потребительских кредитов в 700 млн. рублей эффективность от внедрения предложенного метода составляет 41%. Внедрение разработанного метода рентабельно для банков, среднегодовой портфель потребительских кредитов которых составляет более 496,4 млн. рублей.

Стоит отметить, что вероятность верно классифицированных наблюдений не отражает всех преимуществ разработанного метода, т. к. бинарная классификация физических лиц по кредитоспособности является лишь одним из вариантов использования данного метода. В связи с этим иной положительный эффект от применения разработанного метода, а именно увеличение финансовой стабильности банка в секторе кредитования физических лиц к колебаниям рыночной конъюнктуры оценён в описании следующего положения основных научных результатов работы.

4. Разработана система поддержки принятия решений по установлению процентной ставки на основе классификации физических лиц по кредитоспособности по непрерывной шкале.

В большинстве существующих систем поддержки принятия решений (СППР) некредитоспособным физическим лицам отказывают в выдаче кредита, а остальным, которые подают заявку на один и тот же тип кредита, устанавливают одинаковую процентную ставку. Разработанная СППР включает алгоритм и код в программной среде Wolfram Mathematica 8. На входе в СППР поступает набор характеристик физического лица из заявки на кредит $x^{(j)} = \{x_1^{(j)}, x_2^{(j)}, \dots, x_m^{(j)}\}$. Затем определяется класс кредитоспособности клиента $Q_j = Q(x^{(j)}, w)$. Далее вычисляется процентная ставка по кредиту для клиента $P_j = P(Q_j)$. На выходе СППР предоставляет множество некредитоспособных клиентов, которым следует отказать в выдаче кредита, а для каждого j -го кредитоспособного клиента – значение процентной ставки по кредиту (P_j).

Среди затрат банка потери по ссудам напрямую зависят от действий заёмщиков. Для выражения процентной ставки по кредиту через кредитоспособность достаточно выполнить следующие действия:

1. Составить множество $S = \{c_i, Q_i\}_{i=1}^l$, где $c_i = \frac{L_i}{M_i}$, L_i – потери по i -ой ссуде, c_i – потери по i -ой ссуде на единицу выданных средств, M_i – сумма кредита, а Q_i – класс кредитоспособности клиента взявшего i -ый кредит.
2. Точки множества S аппроксимировать некоторой функцией $c(Q)$, где $c_i = c(Q_i)$, например с помощью полинома степени k .

В результате ожидаемые потери по ссуде можно представить в виде функции $c(Q) = \sum_{i=0}^k a_i Q^i$, $a_i \in R, \forall i = 0, 1, 2, \dots, k$. Следовательно, процентная ставка по кредиту определяется, как функция $P(Q) = P_{min} + \sum_{i=0}^k a_i Q^i$.

Минимальная процентная ставка по кредиту (ставка без учёта кредитного риска) вычисляется, как сумма затрат на обслуживание собственного капитала, на привлечение кредита и на ведение банковской деятельности, делённые на объём выданных кредитов.

$$P_{min} = \frac{S r_s + \sum_{i=1}^z d_i r_i + E}{M_{Total}},$$

где S – собственный капитал, r_s – требуемая норма доходности на собственный капитал без учета ожидаемых потерь по ссудам, d_i – величина i -го депозита, r_i – стоимость i -го депозита, $\sum_{i=1}^z d_i r_i$ – затраты банка на привлечение ресурсов, E – организационные затраты банка, M_{Total} – общая сумма выданных кредитов.

Проведённый анализ риска по методу Монте-Карло показал более высокую финансовую устойчивость коммерческого банка в секторе потребительского кредитования при использовании разработанной СППР, снизив веро-

ятность отзыва лицензии по причине нарушения норматива достаточности собственных средств (Н1.0) на величину от 2,89% до 10,93%.

5. Предложен алгоритм параметрической оценки атрибутов кредита при многопериодном взаимодействии с клиентом на основе классификации физического лица по кредитоспособности.

Алгоритм основан на вычислении значений параметрического уравнения, где в качестве параметра выступает класс кредитоспособности и предназначен для взаимодействия с клиентами по кредитным картам и по кредитным линиям с использованием периодического обновления информации о заёмщике. Например, информация по сети интернет может ежемесячно поступать от работодателя о заработной плате клиента или от товарищества собственников жилья об оплате коммунальных услуг. Отличие от существующих алгоритмов (помимо применения классификации физических лиц по кредитоспособности по непрерывной шкале), где характеристики физического лица и атрибуты предполагаемого кредита анализируются вместе, состоит в осуществлении анализа кредитоспособности физического лица независимо от атрибутов предполагаемой кредитной сделки. Применение подобного алгоритма позволяет построить множество подпространств уровня, каждое из которых выражает комбинации атрибутов кредита с премией за риск, которая соответствует классу кредитоспособности клиента.

Алгоритм ориентирован для использования в следующих ситуациях:

1. обработка новой заявки на кредитную карту или кредитную линию;
2. корректировка атрибутов по действующей кредитной карте или по открытой кредитной линии (обозначим как кредит).

При обработке новой заявки на кредит, на **вход** в СППР поступает частная информация заявителя, некоторого j -го физического лица, представленная в виде набора характеристик $x^{(j)} = (x_1^{(j)}, x_2^{(j)}, \dots, x_m^{(j)})$.

На первом этапе вычисляется значение класса кредитоспособности $Q_j = Q(x^{(j)}, w)$, описание которого и метод вычисления подробно описаны в предыдущих положениях.

На втором этапе значение класса подставляется в параметрическое уравнение $F(P, T, M) = Q$, где функция F характеризует уровень риска кредита в зависимости от основных атрибутов, $F : P \times T \times M \rightarrow [0, 1]$, T – продолжительность кредита, M – сумма кредита, P – процентная ставка по кредиту. Данное параметрическое уравнение описывает множество поверхностей уровня, каждая из которых отображает комбинации условий кредитования с премией за риск, соответствующей фиксированному значению класса кредитоспособности. Например, j -му клиенту с классом кредитоспособности Q_j соответствует поверхность уровня $F(P, T, M) = Q_j$. СППР визуализирует поверхность уровня допустимых вариантов кредитования и вычис-

ляет диапазон варьирования атрибутов кредитной сделки для j -го клиента ($T \in [T_{min}, T_{max}], P \in [P_{min}, P_{max}], M \in [M_{min}, M_{max}]$).

На третьем этапе происходит взаимодействие с клиентом.

На этапе 3 а, клиент выбирает конкретное значение атрибута кредита, например срок кредита ($T = T_j$) из допустимого диапазона $[T_{min}, T_{max}]$. СППР вычисляет линию уровня из оставшихся атрибутов (M и P), подставив фиксированное значение в параметрическое уравнение ($F(P, T_j, M) = Q_j$). Данную процедуру можно представить, как проведение гиперплоскости через выбранное значение T_j параллельно осям P и M . В результате СППР вычисляет диапазон варьирования для оставшихся атрибутов $P \in [P'_{min}, P'_{max}]$, $M \in [M'_{min}, M'_{max}]$.

На этапе 3 б, клиент фиксирует следующий атрибут, например объем кредита ($M = M_j$). СППР вычисляет значение процентной ставки по кредиту $F(P, T_j, M_j) = Q_j$. Более того, СППР вычисляет для дальнейшей работы с клиентом величину кредитного лимита ($M_j^{lim} = F_M^{-1}(Q_j, T_j, P'_{max})$).

В результате на **выходе** СППР предоставляет значение атрибутов кредита для j -го клиента P_j, M_j и T_j , а так же величину кредитного лимита M_j^{lim} .

При работе по действующему кредиту для j -го клиента, на **вход** в СППР поступают значение процентной ставки по кредиту $P_j(t-1)$ и величины кредитного лимита на предыдущем этапе вычислений $M_j^{lim}(t-1)$, срока кредита $T(t-1)$, частная информация j -го заёмщика на текущий момент времени $x^{(j)}(t)$ и значение кредитоспособности в предыдущий момент времени $Q_j(t-1)$.

На первом этапе вычисляется класс кредитоспособности на текущий момент времени ($Q_j(t) = Q(x^{(j)}(t), w)$).

На втором этапе вычисляются следующие значения:

- процентной ставки по кредиту ($P_j(t) = F_P^{-1}(Q_j(t), T_j(t), M_j(t))$, где $T_j(t)$ – срок кредита для j -го клиента на момент t);
- величины кредитного лимита ($M_j^{lim}(t) = F_M^{-1}(Q_j(t), T_j(t), P_{max})$).

В результате на **выходе** СППР предоставляет для банка новую величину кредитного лимита $M_j^{lim}(t)$ и процентную ставку $P_j(t)$ для j -го клиента.

Для того чтобы получить множество поверхностей уровня $F(P, T, M) = Q$, банку достаточно вычислить зависимость ожидаемых затрат по кредиту от атрибутов кредитной сделки при различном классе кредитоспособности клиента. Для этого в осях P , T и M необходимо составить множество $H = \{f_i, Q_i\}_{i=1}^g$, где f_i – потери по кредиту i на единицу выданных средств, а Q_i – класс кредитоспособности заёмщика, взявшего i -ый кредит. Например, если заёмщик взял в банке кредит на сумму M_i и срок T_i , то в точке (f_i, T_i, M_i) отмечается кредитоспособность Q_i . Полученное

множество точек необходимо аппроксимировать некоторой поверхностью, например с помощью следующих полиномов:

- для вычисления процентной ставки по кредиту:

$$P = F_P^{-1}(Q, T, M) = \sum_{i=0}^k (q_i M^i + p_i T^i + h_i Q^i), \quad q_i, p_i, h_i \in R, i = 0, 1, \dots, k,$$

где F_P^{-1} – обратная функция F , взятая по аргументу P ;

- для вычисления кредитного лимита:

$$M^{lim} = F_M^{-1}(Q, T, P) = \sum_{i=0}^k (w_i P^i + u_i T^i + v_i Q^i), \quad w_i, u_i, v_i \in R, i = 0, 1, \dots, k,$$

где F_M^{-1} – обратная функция F , взятая по аргументу M .

Управлять атрибутами по кредиту на основе динамики класса кредитоспособности заёмщиков банк может следующим образом. Допустим, класс кредитоспособности j -го заёмщика за период Δt изменилась с α до β , например, по информации, предоставленной работодателем в банк по интернету. Банк на основе данной информации может принять следующие решения:

1. изменить кредитный лимит клиенту на величину $\Delta M_j^{lim}(t)$:

$$\Delta M_j^{lim}(t) = M_j^{lim}(t) - M_j^{lim}(t-1) = F_M^{-1}(P(t), T(t), \beta) - F_M^{-1}(P(t-1), T(t-1), \alpha),$$

в случае, если прочие параметры неизменны, данное выражение можно привести к следующему виду: $\Delta M_j^{lim}(t) = \sum_{i=0}^k (w_i P^i + u_i T^i + v_i \beta^i) - \sum_{i=0}^k (w_i P^i + u_i T^i + v_i \alpha^i) = \sum_{i=0}^k v_i (\beta^i - \alpha^i)$.

2. изменить процентную ставку по кредиту на величину $\Delta P_j(t)$:

$$\Delta P_j(t) = P_j(t) - P_j(t-1) = F_P^{-1}(M(t), T(t), \beta) - F_P^{-1}(M(t-1), T(t-1), \alpha),$$

при конкретном значении суммы кредита M величина $\Delta P_j(t)$ может выступать, как разность процентных ставок по кредиту при величине кредита M соответствующей разнице класса кредитоспособности: $\Delta P_j(t) = \sum_{i=0}^k (q_i M^i + p_i T^i + h_i \beta^i) - \sum_{i=0}^k (q_i M^i + p_i T^i + h_i \alpha^i) = \sum_{i=0}^k h_i (\beta^i - \alpha^i)$.

Функционирование предложенной системы поддержки принятия решений по предложенному алгоритму продемонстрировано на следующем примере. На Рисунке 1 представлены комбинации условий кредитования для j -го клиента в различных состояниях (точки $A_0 - A_4$):

- Точка A_0 . Клиент взял кредитную карту или открыл кредитную линию с лимитом M_0^{lim} , по которой взял M_0 заёмных средств по процентной ставке $P_0 = F_P^{-1}(Q_j(t_0), M_0)$.
- Точка A_1 . Клиент взял в банке дополнительные заёмные средства до величины кредитного лимита M_0^{lim} , при том же классе кредитоспособности $Q_j(t_1) = Q_j(t_0)$ и процентная ставка составила $P_1 = F_P^{-1}(Q_j(t_1), M_0^{lim})$.

- Точка A_2 . У клиента сгорел дом, что понизило класс кредитоспособности до $Q_j(t_2)$. Процентная ставка по новым кредитам при прежнем кредитном лимите повысилась на величину $\sum_{i=0}^k h_i(Q_j(t_2)^i - Q_j(t_1)^i)$ до уровня $P_2 = F_P^{-1}(Q_j(t_2), M_0^{lim})$.
- Точка A_3 . Клиенту потребовались дополнительные займы. Однако клиент уже взял самый дорогой кредит при своём классе кредитоспособности. Для получения дополнительных средств клиент оставил свой автомобиль под залог в банк. Банк увеличил кредитный лимит на величину $\sum_{i=0}^k v_i(Q_j(t_3)^i - Q_j(t_2)^i)$, который составил $M_3^{lim} = F_M^{-1}(P_3, Q_j(t_3))$. Заёмщику потребовались все доступные заёмные ресурсы, поэтому процентная ставка осталась наиболее высокой $P_3 = P_2$.
- Точка A_4 . Клиент на рассматриваемый момент не имеет задолженности перед банком, класс кредитоспособности остался прежним $Q_j(t_4) = Q_j(t_3)$. Однако клиенту потребовалось взять кредит величиной M_0 , которой теперь соответствует процентная ставка $P_4 = F_P^{-1}(Q_j(t_4), M_0)$.

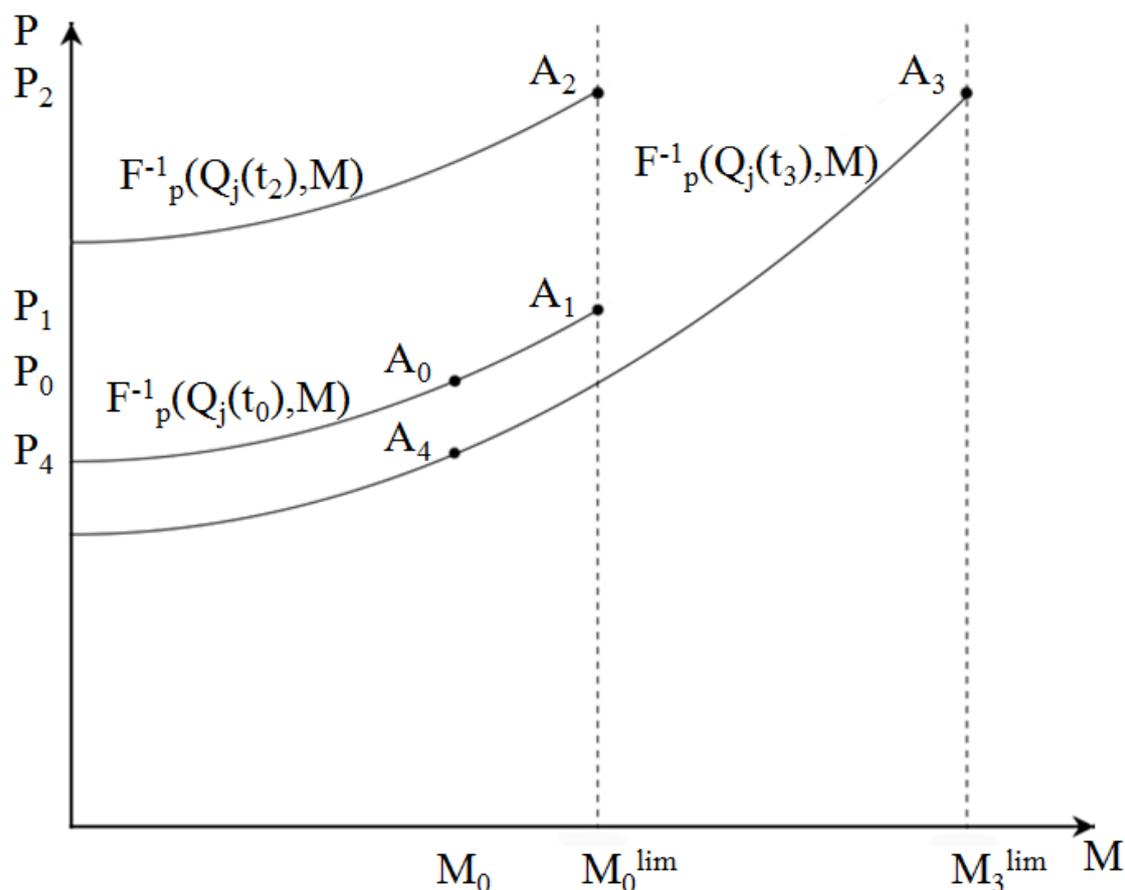


Рис. 1: Комбинации условий кредитования

Применение разработанного алгоритма обеспечит банку возможность визуализировать допустимые варианты кредитования клиентам в соответствии с классом кредитоспособности и управлять атрибутами кредитного

контракта по динамике данного класса. Оперативное управление данными атрибутами обеспечит снижение кредитного риска.

Более того, банк безразличен к выбору клиента комбинаций атрибутов кредита, соответствующих множеству значений поверхности уровня ($F(P,T,M) = Q_j$), потому что премия за риск отражена соответствующим образом в процентной ставке по кредиту. Таким образом, банк может предоставить право выбора условий кредитования клиенту из указанного множества.

В таком случае клиент выберет наиболее удобный для себя вариант выплаты кредита. Удобство клиента по кредитным выплатам так же обеспечит снижение кредитного риска. Стоит отметить, что выбирая вариант кредитования, клиент максимизирует свою функцию полезности. В результате, предоставление подобной возможности может выступать в качестве фактора при поиске кредитора.

Публикации по теме диссертации

1. Романюк К. А. Концепция метода оценки кредитоспособности физических лиц / К. А. Романюк // Финансы и кредит. – 2015. – №24. – С. 45-53. – 0,8 п.л.
2. Романюк К. А. Метод оценки кредитоспособности физических лиц по непрерывной шкале / К. А. Романюк // Экономические науки. – 2015. – № 125. – С. 109-116. – 0,6 п.л.
3. Романюк К. А. Ценовая стратегия банка, основанная на непрерывном отображении частной информации физических лиц в процентной ставке по кредиту / К. А. Романюк // Российское предпринимательство. – 2015. – Т. 16, № 14. – С. 2157-2174. – 0,9 п.л.
4. Романюк К. А. Построение валюты минимального риска из основных мировых валют / К. А. Романюк // Управление экономическими системами: электронный научный журнал. – 2013. – №49. – 0,7 п.л.
5. Romanyuk K. Mortgage lending for slum clearance / K. Romanyuk // Procedia Engineering. – 2015. – № 117. – С. 304-308. – 0,5 п.л.
6. Romanyuk K. Concept of a decision support system in a loan granting based on a continuous price function / K. Romanyuk // Conference on Intelligent Systems, London, 2015. – IEEE, 2015. – С. 105-111. – 0,8 п.л.
7. Romanyuk K. Credit scoring based on a continuous scale for on-line credit quality control / K. Romanyuk // Conference on Evolving and Adaptive Intelligent Systems, Douai, 2015. – IEEE, 2015. – С. 158-162. – 0,6 п.л.