


Федеральное государственное автономное  
образовательное учреждение высшего образования  
“Санкт-Петербургский политехнический университет Петра Великого”

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»

Диссертация на соискание учёной степени  
кандидата экономических наук

Научный руководитель:  
доктор экономических наук,  
профессор Ильин Игорь Васильевич

Санкт-Петербург — 2016

# Оглавление

Введение . . . . .	3
<b>1 Рынок кредитования физических лиц . . . . .</b>	<b>10</b>
1.1 Характеристика рынка кредитования физических лиц России . . . . .	10
1.2 Теоретические аспекты кредитования физических лиц . . . . .	21
1.3 Практика классификации физических лиц . . . . .	36
1.4 Обзор методов классификации объектов . . . . .	48
<b>2 Разработка метода классификации объектов по непрерывной шкале . . . . .</b>	<b>61</b>
2.1 Модель классификации объектов по непрерывной шкале . . . . .	61
2.2 Модифицированный метод рандомизированных сводных показателей . . . . .	66
2.3 Метод классификации объектов по непрерывной шкале . . . . .	81
<b>3 Применение метода классификации объектов по непрерывной шкале при кредитовании физических лиц . . . . .</b>	<b>93</b>
3.1 Классификация физических лиц по базе кредитных историй . . . . .	93
3.2 Система поддержки принятия решений по обоснованию процентной ставки физическим лицам . . . . .	106
3.3 Экономический эффект и эффективность . . . . .	122
3.4 Особенности ценообразования кредита для физических лиц . . . . .	132
Заключение . . . . .	146
Список литературы . . . . .	148
<b>А Вычисления класса кредитоспособности . . . . .</b>	<b>164</b>

## Введение

**Актуальность темы диссертационной работы.** За последние два года, после введения рядом стран ограничительных мер в отношении российской экономики, и снижения 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 шка-

ле. 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 рассмотрен вопрос классификации объектов по непрерывной шкале, примененная 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. Наем, 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. Вуцичевич, 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, 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 того чтобы выявить подходящую модель 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 отклонения процентной ставки по кредиту от потерь по кредиту, увеличивая финансовую стабильность к колебаниям рыночной конъюнктуры; возможность визуализировать допустимые комбинации условий кредита в соответствии с классом кредитоспособности физического лица и корректировать данные условия по динамике, вычисленного класса; снижение потерь по кредитам на единицу выданных займов.

**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 которым результаты работы также 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.



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 // IEEE Conference on Evolving and Adaptive Intelligent Systems, Douai, 2015. – IEEE, 2015. – С. 158-162. – 0,6 п.л.

**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.

Благодарность Елманову Николаю Александровичу за то, что в годы моего студенчества сориентировал меня в область кредитования физических лиц, а так же за ценные комментарии в период написания диссертации.

## Глава 1.

### Рынок кредитования физических лиц

В данной главе рассмотрена текущая ситуация на рынке кредитования физических лиц. Проанализированы теоретические аспекты асимметрии информации и её воздействие на рынок кредитования физических лиц. Проанализирована значимость классификации клиентов по кредитоспособности для мировой экономики и существующие подходы по данной классификации. Приведено теоретическое обоснование и банковская практика дифференциации процентной ставки по кредиту.

#### 1.1 Характеристика рынка кредитования физических лиц России

В последние годы существенно выросло число отозванных лицензий Центральным Банком у коммерческих банков. В Таблице 1.1 представлено число отозванных лицензий у банков без учета случаев ликвидации юридического лица (возможно по причине слияния, поглощения и т.п.) за период с 2012 по 2015. За 2014 и 2015 год число отозванных лицензий у банков по сравнению с 2012 и 2013 годами увеличилось более чем в 3,5 раза. В 2014 году рядом стран были приняты ограничительные меры в отношении российской экономики, а так же было снижение финансовых поступлений в страну после падения цен на нефть, что продолжилось в 2015 году. Таким образом банковская система Российской Федерации обладает низкой стабильностью к колебаниям рыночной конъюнктуры.

Сектор потребительского кредитования является весьма нестабильным. За 2011 и 2012 годы рассматриваемый сектор вырос на 35,9% и 39,4% соответствен-

но, что многократно превысило развитие по другим направлениям банковской деятельности. В частности, сектор корпоративного кредитования за 2011 и 2012 годы вырос на 26% и 12,7% соответственно [19]. Доля активов банков в секторе потребительского кредитования существенно выросла с 13,3% в 2012 г. до 17,3% в 2014 г., а в 2015 г. наблюдалось стремительное снижение данной доли до 14,6% [1]. Однако снижение носит лишь относительный характер, потому что совокупные активы банковского сектора существенно увеличились с 57,42 трлн. рублей (01.01.2014) до 77,65 трлн. рублей (01.01.2015).

Таблица 1.1: Число отозванных лицензий без учёта случаев ликвидации юридического лица

Год	Число
2012	22
2013	32
2014	88
2015	103

Стоит отметить, что объемы кредитования населения за первую половину 2015 г. падают с 11,3 трлн. рублей (01.01.2015) до 10,7 трлн. рублей (01.07.2015), в тоже время просроченные обязательства растут с 667,5 млрд. рублей (01.01.2015) до 806,3 млрд. рублей (01.07.2015) [1]. Если взять более широкий временной диапазон, то на 01.01.2014 величина просроченных обязательств составила 440,3 млрд. рублей. В процентном отношении к объему выданных кредитов просроченные обязательства с 4,42% к 01.01.2014 выросли до 5,89% к 01.01.2015 и до 7,45% к 01.07.2015. Таким образом увеличение доли просроченных обязательств в секторе кредитования физических лиц за полтора года составило почти 60%.

В общем, целенаправленное улучшение характеристик коммерческих банков: уменьшение доли просроченных платежей, безнадежных кредитов или увеличение объема рынка даже на 1%, при прочих равных условиях, принесет огромные прибыли для коммерческих банков, что является одной из причин активной разработки исследователями новых методов классификации физических лиц по кредитоспособности (кредитного скоринга) [31,51]. В условиях текущей негативной динамики увеличения просроченной задолженности в секторе потребительского кредитования, а так же увеличения количества отозванных

лицензий у банков, разработка методов и инструментов, для более качественного управления кредитным риском в рассматриваемом секторе и повышения финансовой стабильности банка, является актуальным направлением исследования.

Непредсказуемые колебания просроченной задолженности несут угрозу для финансовой устойчивости коммерческих банков [6]. В 2013 году вступили в силу изменения ЦБ РФ в отношении формирования банками резервов под выдаваемые кредиты. В частности, мера, которая начала действовать с 1 марта 2013 года - это "повышение ставок резервирования по необеспеченным розничным кредитам, сгруппированным в портфели однородных ссуд" [18]. Норма увеличилась в два раза: до вступления изменений в силу банкам необходимо было резервировать 1% от размера портфеля непросроченных потребительских кредитов, после 2%, с просрочкой до 30 дней, раньше 3%, после 6% и т.д. А по кредитам, по которым просрочка больше года, теперь придется начислять 100-процентные резервы вместо 75% [17]. При данном увеличении нормативов наличие просроченной задолженности со стороны категории заёмщиков, которая указана выше, становится более дорогой для коммерческих банков. Таким образом, ЦБ РФ сделал наличие просроченной задолженности более дорогой для банков с целью увеличения качества управления кредитным риском. В 2014 году ЦБ РФ продолжал мониторинг сектора кредитования физических лиц с целью идентификации негативных тенденций в данном секторе с выделением конкретных банков, действия которых порождают данные тенденции [21].

Одной из целей ЦБ РФ является "обеспечение стабильности финансового рынка Российской Федерации" [27]. Чем менее стабильна банковская система, тем труднее банку получить вклады от клиентов, что повышает процентную ставку по кредиту и замедляет экономическую активность в стране. По структуре вкладов населения с 01.01.2014 по 01.07.2015 относительная величина долгосрочных вкладов (более 1 года) снизилась с 61,8% до 43,6%, а относительная величина вкладов от 31 дня до 1 года увеличилась с 19,2% до 40,7%, что может свидетельствовать о снижающемся доверии населения к банкам [1]. Для стабильного развития сектора потребительского кредитования необходима гибкая система управления кредитным риском (система поддержки принятия решений при кредитовании), чем не являются существующие системы, где отсеиваются

некредитоспособные клиенты, а кредит выдается остальным по единой процентной ставке.

Процедура обработки заявки на кредит физическим лицам в России, как правило, проходит в два этапа. На первом этапе банк определяет целесообразность работы с потенциальным заемщиком. Решается вопрос: “отклонить” или “принять” заявку? На втором этапе определяется максимальный объем средств, который будет предоставлен потенциальному заемщику. Запрос может быть удовлетворен в полном объеме, а может быть выполнен частично, что зависит от ожидаемых потерь при выдаче кредита. Ограниченность методов, классификации физических лиц по кредитоспособности, которая не позволяет дифференцировать процентную ставку, приводит к тому, что в некоторых случаях заемщик не получает необходимый объем средств, а банк упускает выгоду, выдавая кредит в меньшем объеме. При гибком ценообразовании кредита существует процентная ставка, при которой риск банка оплачен, а потребность заемщика в кредите удовлетворена в большей степени. Более того, если банк установил двум клиентам одинаковый кредитный лимит из-за одинаковой оценки кредитного риска (одинакового класса кредитоспособности) и первый клиент взял меньше кредитных средств, чем второй, то кредитный риск на единицу выданных средств для второго клиента выше.

На Рисунке 1.1 представлена схема определения процентной ставки для физических лиц. В отличие от юридических лиц, кредитные истории и частная информация физических лиц является конфиденциальной, поэтому в качестве исходных данных выступает информация, которую предоставил заемщик. Данная информация анализируется методами классификации физических лиц по кредитоспособности, большинство из которых осуществляют бинарную классификацию клиентов: “кредитоспособен” и “некредитоспособен”. На основе бинарной классификации банк может принять только два решения о выдаче кредита. Некредитоспособным клиентам кредит выдавать не следует, как обладателям наибольшего кредитного риска. Стоит отметить, что кредитоспособность является мерой, которая связана с вероятностью банкротства заемщика ( $P_b(x)$ ), обладающего набором индивидуальных характеристик ( $x = (x_1, x_2, \dots, x_n)$ ) некоторой убывающей функцией от кредитоспособности ( $P_b(x) = F(Q(x, w))$ ), где  $Q(x, w)$  – кредитоспособность, а  $w = (w_1, w_2, \dots, w_n)$  – вектор параметров). Кре-

дитоспособные клиенты неразличимы между собой, так как принадлежат одному классу. Таким образом вполне обоснованно, что банки используют подобные методы для отсева наименее кредитоспособных клиентов и выдаче кредита остальным по единой процентной ставке. Важно отметить, что подобное ценовая стратегия имеет обратную связь на действия клиентов, в частности порождает неблагоприятный отбор.

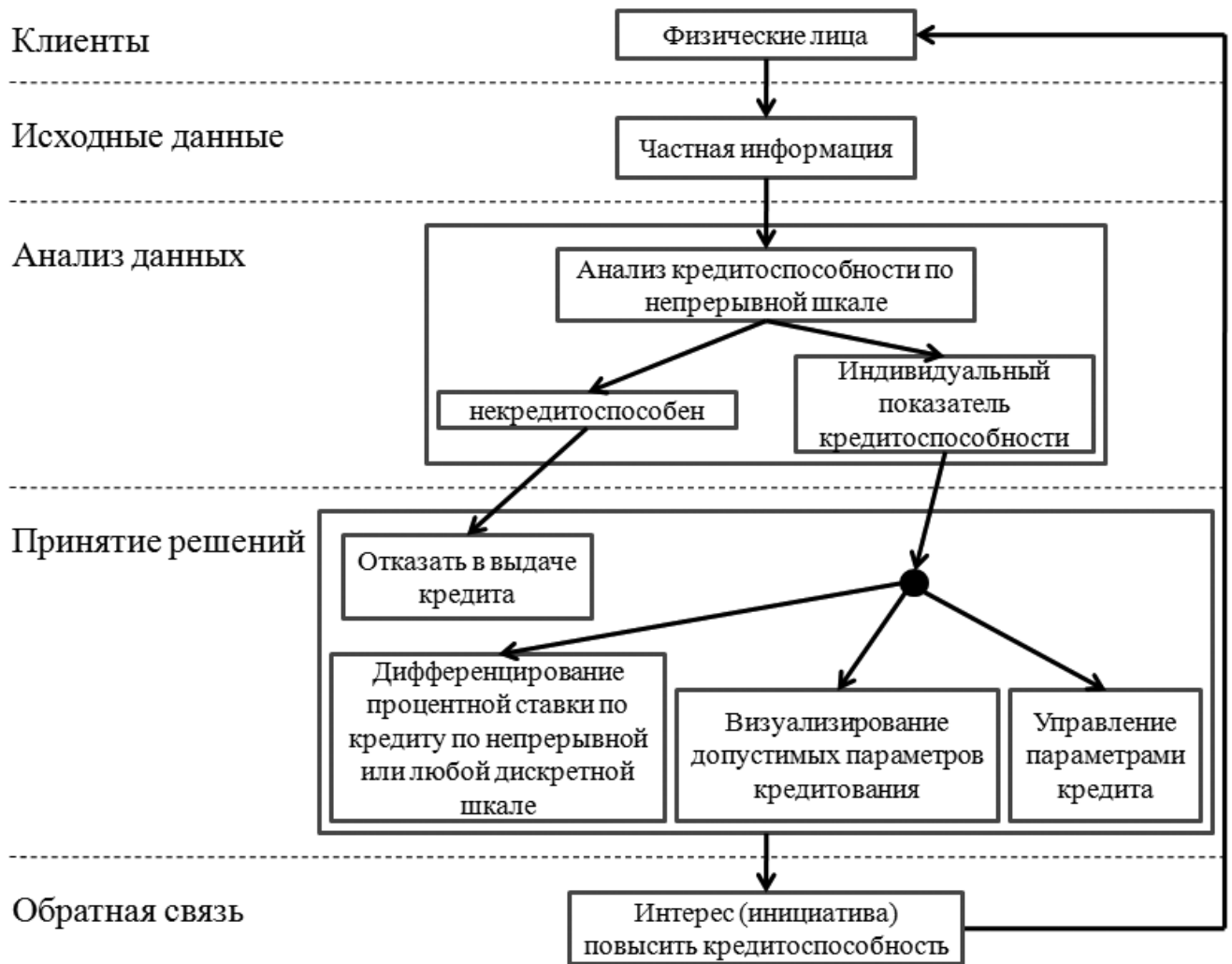


Рисунок 1.1: Схема установления процентной ставки по кредиту физическим лицам

Заемщики имеют различный уровень кредитного риска, что влечет различную стоимость кредита. Стоимость кредита для заемщика с высоким уровнем кредитного риска выше, чем стоимость кредита для заемщика с низким уровнем риска. В большинстве российских банков заемщики платят не за свой индивидуальный уровень риска, а за средний уровень по банку. В данной ситуации возникает один из эффектов асимметрии информации, а именно неблагоприятный отбор клиентов. Асимметрия информации заключается в том, что про-

давец (потенциальный заёмщик) знает качество товара (уровень кредитоспособности клиента), который он продает, а покупатель (банк) не знает точного значения качества, и ориентируется по среднему уровню рынка. Качественный товар стоит дороже, чем некачественный, но так как покупатель знает лишь средний уровень, то он готов предложить среднюю цену. Продавцу с некачественным товаром данная цена выгодна, так как ему переплачивают (платят больше действительной стоимости) за товар, но для продавца с качественным товаром возникает обратная ситуация - ему не доплачивают (платят меньше действительной стоимости) за товар (Рис. 1.2).

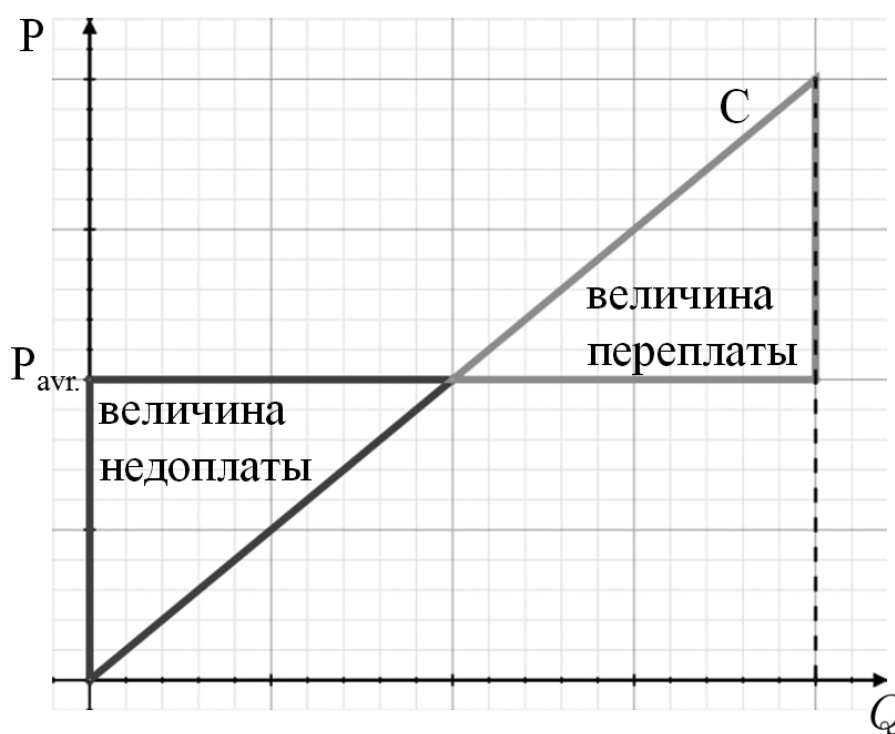


Рисунок 1.2: Несоответствие цены товара затратам на производство при асимметрии информации

Продавец качественного товара предпочтет уйти с данного рынка. Если предположить, что все продавцы с товаром по качеству выше среднего уйдут с данного рынка, то изменится и цена на рынке. Снова возникнут те продавцы, которым переплачивают, и те, кому недоплачивают (Рис. 1.3). Продавцы качественного товара относительно нового среднего уровня предпочтут уйти с рынка, так как им недоплачивают. В результате цена снова снизится и так будет продолжаться до тех пор, пока не останутся лишь продавцы самого некачественного товара.

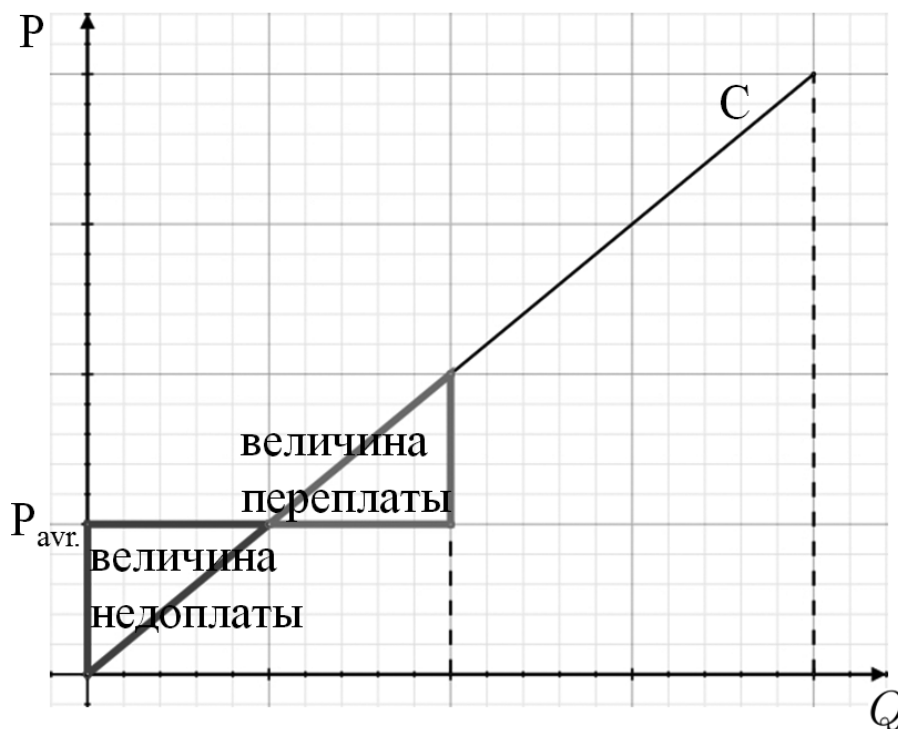


Рисунок 1.3: Снижение цены при уходе с рынка продавцов качественного товара при асимметрии информации

Для соответствия рынка потребительского кредитования, приведенной выше схеме действия асимметрии информации в роли продавца качественного товара должен выступать заёмщик с высокой кредитоспособностью, а в роли покупателя банк. Подобное представление заёмщика в качестве продавца может показаться подгонкой под ситуацию асимметрии информации. Однако кредитование можно представить различным образом. С одной стороны, банк продаёт текущие деньги в обмен на будущие деньги с процентными платежами. С другой стороны, заёмщик продаёт будущие деньги в обмен на текущие деньги с дисконтом. В результате представление заёмщика в качестве продавца является оправданным.

Большинство российских банков устанавливают единую, среднюю процентную ставку (“house rate”) для физических лиц, которые подают заявку на один тип кредита, как во многих западных банках [88]. Получается, что для заёмщика с низким уровнем кредитного риска средняя процентная ставка выше стоимости кредита, которая соответствует данному уровню. Напротив, для заёмщика с высоким уровнем кредитного риска средняя процентная ставка ниже стоимости кредита. Заемщики с низким уровнем риска материально заинтере-



ресованы не использовать кредит в банках с единой усредненной процентной ставкой, а перейти в организацию, где при прочих равных условиях производится дифференциация стоимости кредита по риску. Однако заемщики с высоким уровнем кредитного риска заинтересованы в том, чтобы брать кредит в банках с единой усредненной процентной ставкой. В итоге, банк, устанавливающий усредненную процентную ставку, остается с клиентами, которые обладают высоким уровнем кредитного риска.

Для того чтобы дифференцировать процентную ставку для заемщиков, коммерческим банкам несмотря на значимость использования подходящих вычислительных и информационных систем, необходимо классифицировать физических лиц по кредитоспособности по шкале с числом градаций, как минимум больше чем 2. Результат подобной классификации может быть использован, чтобы в автоматизированном режиме дифференцировать условия кредитования для заемщика. Существует понятие супервосстановление, которое характеризует способность человека после получения нагрузки и соответствующего восстановления быть готовым справляться с более высокой нагрузкой. Доступность кредита влияет на экономическую активность [71]. Таким образом, нынешняя банковская система создаёт барьер для развития физических лиц с высокой кредитоспособностью.

Независимость процентной ставки по кредиту для физических лиц от характеристик, по которым определяется кредитоспособность в большинстве банков имеет негативный эффект, как на физических лиц, так и на банки. У физических лиц снижается спрос на товары и услуги из-за недостатка ресурсов, что препятствует оптимизации потребления и реализации в полной мере своего развития. Для банков усредненное ценообразование процентной ставки по кредиту снижает конкурентоспособность и долю платежеспособных клиентов, что увеличивает уровень риска кредитного портфеля, и уменьшает его объем. В результате эффективность использования ресурсов падает. Кроме того, заемщик в финансовом отношении не ощущает позитивного эффекта от предоставления частной информации в банк. Взаимосвязь между предоставляемой информацией и процентной ставкой банк может использовать для управления составом заемщиков. Предоставляя кредит конкретной категории заемщиков на особых

условиях, более выгодных, чем остальным, можно привлечь данную категорию со всего рынка.

В секторе корпоративного кредитования применяется принцип соответствия процентной ставки уровню кредитного риска (*risk-based pricing*). В связи с этим, политика банков в секторе потребительского кредитования выглядит непоследовательной. С одной стороны, для снижения кредитного риска банки собирают и анализируют информацию о клиентах при принятии решения о выдаче кредита. С другой стороны, большинство банков не используют проанализированную информацию для дифференциации процентной ставки. После финансового кризиса 2008 года, в основе которого было обрушение рынка ипотечного кредитования многие эксперты осознали, что необходимо учитывать кредитоспособность при выдаче кредита. В последние годы ведутся работы по реализации принципа соответствия процентной ставки уровню кредитного риска при ипотечном кредитовании [106].

В российской банковской практике при кредитовании физических лиц существует пример дифференцирования процентной ставки по кредиту в соответствии с предоставляемой частной информацией. В публичном акционерном обществе «Совкомбанк» в 2014 году была проведена дифференциация условий кредитования, которые сопровождались широкомасштабной рекламной кампанией. Текст оферты содержал два сообщения:

1. «Вы пенсионер? Дадим Вам 12%»;
2. «Ваш стаж 20 лет? Дадим Вам 12%».

Данная оферта означала, что пенсионерам и лицам с трудовым стажем 20 лет и более в указанном банке предоставляется кредит под процентную ставку 12% годовых<sup>1</sup>. Средневзвешенные процентные ставки по предоставленным кредитам физическим лицам в рублях в 2014 году по 30 крупнейшим банкам, принимали следующие значения (Табл.1.2):

- на срок от 91 до 180 дней процентные ставки колебались от 27,48% до 34,93%, а среднее значение по месяцам составило 31,72%;

---

<sup>1</sup>Указанная процентная ставка действовала только для кредита в 100 тысяч рублей на полгода.

– на срок от 181 дня до 1 года процентные ставки колебались от 18,49% до 27,78%, а среднее значение по месяцам составило 23,57%.

Таблица 1.2: Средневзвешенные процентные ставки по кредитам, предоставленным кредитными организациями физическим лицам в рублях в 2014 году по 30 крупнейшим банкам. (Источник Банк России)

	от 90 до 180 дней	от 181 дня до 1 года
Январь	34,93%	22,37%
Февраль	32,72%	22,01%
Март	32,25%	24,81%
Апрель	31,62%	18,49%
Май	34,56%	24,46%
Июнь	31,20%	23,66%
Июль	28,59%	22,77%
Август	31,61%	23,10%
Сентябрь	27,48%	23,96%
Октябрь	32,24%	24,54%
Ноябрь	31,16%	24,88%
Декабрь	32,28%	27,78%

Кредит на полгода формально попадает в категорию от 181 дня до 1 года, но весьма близок к категории от 91 до 180 дней. Таким образом, кредит, предоставленный Совкомбанком для указанных в оферте категорий, почти в 2 раза дешевле, чем по рынку в категории до 1 года и более чем в 2,5 раза в категории до полугода. Таким образом, заёмщики ПАО «Совкомбанка» с трудовым стажем более 20 лет или являющиеся пенсионерами были заинтересованы предоставить частную информацию в банк, чтобы получить более дешёвый кредит. Информация, полученная банком по данной оферте, ограничивалась подтверждением статуса пенсионера. В 2015 году «Совкомбанк» повторил подобную оферту, что косвенно свидетельствует о её доходности. Подобный опыт стимулирования физических лиц предоставлять частную информацию в банк целесообразно распространить на всю информацию, по которой определяется кредитоспособность клиента.

Как определить взаимосвязь результата классификация физических лиц по кредитоспособности и процентной ставки по кредиту? При управлении предприятием существует система определения покрытия затрат, называемая direct-

costing, основанная на выделении переменных затрат и их учёте при определении покрытия расходов [26]. Для того чтобы дифференцировать процентную ставку в соответствии с классом кредитоспособности (кредитным риском) клиента, в структуре затрат банка следует выделить затраты, которые напрямую зависят от действий заёмщиков. По степени влияния клиента в потребительском кредитовании на затраты банка можно выделить:

- потери по ссудам, которые напрямую зависят от действий заёмщика;
- базовые затраты – затраты на привлечение ресурсов, аренда помещений для ведения банковской деятельности, заработная плата сотрудникам банка и т.д., которые не зависят от отдельно взятого заёмщика.

Чем больше кредитный риск, тем выше себестоимость кредита [64]. В структуре затрат банка данное положение соответствует тому, что чем выше кредитный риск (ниже кредитоспособность), тем выше потери по ссудам. Однако фактические потери по ссудам известны только после окончательного исполнения контракта по кредиту. На момент выдачи кредита банк может оценивать потери по ссудам. Стоит отметить, что кредитоспособность физического лица может принимать бесконечное число значений. Для адекватной классификации физических лиц по кредитоспособности необходимо применение непрерывной шкалы. Каждый клиент по данной шкале получит индивидуальный класс кредитоспособности. По индивидуальному уровню кредитоспособности может быть определено однозначное (биективное) отображение на процентную ставку, то есть установлена индивидуальная процентная ставка. Дифференцирование процентной ставки в соответствии с себестоимостью, которая зависит от ожидаемых потерь по выданной ссуде, обеспечивает два основных преимущества для банка:

1. Возможность привлечь надёжных клиентов. Клиенты с высоким уровнем кредитоспособности больше остальных заинтересованы в переходе в банк, где дифференцируется процентная ставка в соответствии с ожидаемыми потерями по ссудам, потому что при усреднении процентной ставки данная категория клиентов переплачивает больше всего. Данное преимущество будет подробно разобрано в следующем параграфе, в котором

будет смоделировано рыночное поведение клиентов банка в секторе потребительского кредитования при дифференциации процентной ставки в соответствии с классом кредитоспособности.

2. Стимулирование клиентов предоставлять частную информацию в банк. Данный вопрос будет разобран более подробно в параграфе 3.4.

Стоит отметить, что если банк дифференцирует кредитный лимит для клиентов не дифференцируя процентную ставку, то принцип соответствия процентной ставки уровню кредитного риска (*risk-based pricing*) не выполняется. Например, как было упомянуто ранее, два клиента с одинаковой кредитоспособностью получают одинаковый кредитный лимит в банке. Первый клиент берёт больше кредитных средств, чем второй. Процентная ставка одинакова. Однако величина риска на единицу выданных средств для первого клиента выше. Таким образом, на рынке потребительского кредитования необходимы новые системы поддержки принятия решений для объективного дифференцирования условий кредитования в соответствии с кредитным риском клиента.

## **1.2 Теоретические аспекты кредитования физических лиц**

В текущем параграфе рассмотрены теоретические аспекты, связанные с асимметричным распределением информации на рынке кредитования физических лиц. Смоделирована динамика состава заёмщиков банков при различной дифференциации процентной ставки по кредиту для определения эффекта неблагоприятного отбора.

Асимметрия информации - ситуация, в которой продавец и покупатель обладают различной информацией о качестве товара. На рынке кредитования физических лиц кредитор не знает потенциального заёмщика, его возможностей по возвращению кредита, что лежит в основе асимметрии информации. Использование грубых методов классификации физических лиц по кредитоспособности способствует снижению подобной асимметрии. Асимметрия информации имеет негативные следствия для рынка, такие как неблагоприятный отбор и мораль-

ный риск [3]. В текущей работе больше внимания уделено неблагоприятному отбору.

Неблагоприятный отбор приводит к ухудшению качества товара на рынке [34]. На рынке кредитования физических лиц неблагоприятный отбор действует в форме замещения заёмщиков с низким кредитным риском (высокой кредитоспособностью) заёмщиками с высоким кредитным риском (низкой кредитоспособностью), что увеличивает риск кредитного портфеля. Более того, асимметрия информации увеличивает цену привлечения капитала для фирмы [142]. Если экстраполировать данный результат на рынок кредитования физических лиц, то получится, что для указанной категории клиентов асимметрия информации увеличивает процентную ставку.

Для снижения асимметрии информации на рынке кредитования предназначены кредитные рейтинги (оценка кредитного риска, оценка кредитоспособности) [95]. Кредитные агентства снижая влияют на доступность кредитов фирмам, что оказывает воздействия на экономические показатели фирмы [131]. Допустим, что кредитный рейтинг стал более точным за счёт добавления дополнительных градаций. Фирме, которая в результате данной модернизации получила более высокий рейтинг, банк с гибким ценообразованием предоставит кредит по более низкой процентной ставке, что положительно отразится на эффективности предприятия. Однако результат бинарной классификации весьма ограниченно снижает асимметрию информации.

Информационная инфраструктура является основой при управлении организацией [9]. Информация о заёмщике имеет ценность для банка при принятии управленческих решений. Допустим, банк знает своих наиболее доходных клиентов, а конкурирующий банк данной информацией не обладает. В работе Стиглица и Вайса [128], отмечено, что в данном случае при попытке конкурирующего банка переманить наиболее доходных клиентов более низкими процентными ставками исходный банк в ответ снизит процентные ставки для того, чтобы не допустить перехода. Получается, что банку выгодно не предоставлять информацию о некоторых своих клиентах сторонним организациям, чтобы получать сверхприбыль за счёт монопольного владения информацией о заёмщике. Заёмщику с высокой кредитоспособностью, напротив выгодно предоставлять частную информацию в банки, чтобы получить более дешёвый кредит.

Информационная прозрачность в банковском секторе важна и для экономики страны. Высокий уровень обмена информацией между банками делает кредиты более доступными для фирм [46], что можно использовать при кредитовании физических лиц. Более того, согласно исследованию по кредитным организациям с 1975 по 2006 год, высокий уровень обмена информацией между банками снижает вероятность банкротства банка, причём более отчётливый эффект наблюдается в странах с более низкими доходами [48]. Кредитование более развито и кредитный риск ниже в тех странах, где банки обладают информацией о заёмщиках, причём независимо от источника информации [87]. Обмен информацией между банками о клиентах может быть представлен, как дисциплинарный механизм, стимулирующий заёмщиков в полной мере выполнять условия по кредиту [117]. Таким образом для развития экономики страны важно сделать рынок кредитования физических лиц более прозрачным.

Для уменьшения негативных эффектов от ситуации асимметрии информации Ротшильд и Стиглиц предложили использовать “метод отбора” (screening device) и “самоотбор” (self-selection) [122]. Метод отбора (screening device) — это контракт, условия которого установлены таким образом, чтобы привлечь конкретную категорию клиентов. Принцип самоотбора заключается в том, что клиент, принадлежащий к определённой категории, заинтересован в некоторых условиях контракта больше, чем в других. Понятия метод отбора и самоотбор дополняют друг друга. Метод отбора используют, чтобы клиент осуществил самоотбор. Таким образом, метод отбора и самоотбор используют, чтобы установить взаимосвязь между характеристиками клиента и условиями контракта. Метод отбора изначально был предложен страховым фирмам для дифференциации клиентов и условий страхования.

Согласно однопериодной модели кредитования фирмы, дифференциация процентной ставки является не лучшим методом отбора, потому что более высокая процентная ставка побуждает фирмы браться за более доходные проекты с низкой вероятностью успеха [128]. Однако в соответствии с многопериодной моделью кредитования фирм, банку выгодно предоставлять “хорошим” клиентам<sup>2</sup> кредит под более низкую процентную ставку [128]. Подобный эффект связан с тем, что метод отбора в классическом варианте подразумевает на-

---

<sup>2</sup>клиентам, которые ранее выполняли условия кредита в данном банке

личие асимметрии информации: банк ничего не знает о клиенте. В случае с многопериодной моделью, асимметрия информации снижена за счёт кредитной истории, которая накоплена в банке. Стоит отметить, что в последнее время решения, принимаемые банками о выдаче кредита, стали более информационно обоснованными [114]. При наличии методов классификация физических лиц по кредитоспособности с числом классов более двух поощрять клиентов с высокой кредитоспособностью более низкими процентными ставками можно будет уже при первом обращении за кредитом.

Стоит упомянуть, что на рынке кредитования помимо неблагоприятного отбора банки сталкиваются и с моральным риском. Моральный риск – это одна из ситуаций асимметричного распределения информации, которая заключается в том, что агент (заёмщик) после заключения контракта (взятия кредита) начинает действовать не в интересах принципала (банка). В работе [101] представлена модель, согласно которой вероятность возникновения ситуации морального риска монотонно зависит от ставки по кредиту. Чем выше процентная ставка, тем выше вероятность морального риска. Данный результат получен в предположении, что клиенты неразличимы. В параграфе 3.4 рассматриваются особенности ценообразования кредита с использованием разработанного в текущей работе метода для стимулирования долгосрочных отношений между банком и заёмщиком (физическим лицом), препятствуя неблагоприятного отбора и морального риска в секторе потребительского кредитования.

Наряду с методом самоотбора для уменьшения асимметрии информации существует понятие сигнализирования (signaling), предложенное лауреатом нобелевской премии Майклом Спенсом [127]. В качестве сигнализирования (signaling) Спенс использовал образование для классификации потенциального работника по трудоспособности с использованием бинарной шкалы: трудоспособный работник, нетрудоспособный работник. Аналогично в банковской практике физические лица классифицируются на основе доступной информации по бинарной шкале.

Практические результаты оценки влияния асимметрии информации на рынок потребительского кредитования можно найти в работе [90], в которой посчитано, что неблагоприятный отбор встречается чаще среди женщин, а мо-



ральный риск среди мужчин, а в совокупности данные явления объясняют 20% невыплат по кредитам.

Приведём пример воздействие упомянутой асимметрии информации от использования грубых методов классификации физических лиц по кредитоспособности. Основная идея заключается в том, что количество классов при классификации физических лиц по кредитоспособности влияет на определение условий кредитования и на рыночное поведение клиентов. Далее смоделирована конкуренция банков, использующих различные шкалы при классификации клиентов.

Предпосылки модели. В секторе потребительского кредитования действуют 4 банка. Себестоимость кредита состоит из базовых затрат (затрат на привлечение капитала, на ведение банковской деятельности и т. п.) и потерь по ссудам. Банки имеют одинаковые базовые затраты. Потери по ссудам зависят от действий заёмщиков по обслуживанию кредита. Фактические потери по ссудам известны после окончательного исполнения контракта по кредиту. На момент выдачи кредита банк оценивает потери по ссудам. Чем выше кредитоспособность потенциального заёмщика, тем меньше ожидаемые потери по ссудам. Банки дифференцируют процентную ставку в соответствии с возможностями используемого метода классификации физических лиц по кредитоспособности. Для большей наглядности банки при классификации физических лиц по кредитоспособности используют одинаковый уровень допустимой кредитоспособности ( $Q_0$ ).

Банк № 1 выдаёт кредит всем клиентам, не классифицируя их кредитоспособность. Процентная ставка одинакова для всех клиентов (1.1), и равна средней стоимости всех запрошенных кредитов ( $P_1$ ).

$$P_1(Q_i) = P_1, \forall Q_i \quad (1.1)$$

Банк № 2 классифицирует физических лиц по кредитоспособности с использованием наиболее распространённой, бинарной шкалы: “кредитоспособен”, “некредитоспособен”. Кредит выдаётся только “кредитоспособным” клиентам (1.2). Процентная ставка одинакова для “кредитоспособных” клиентов, то есть для клиентов, чья действительная кредитоспособность выше допустимого уровня ( $Q_0$ ). Процентная ставка равна средней стоимости всех выданных кредитов

( $P_2$ ). Данный подход к принятию решений о выдаче кредита более надёжный для банка, так как кредит не выдаётся клиентам с очень низким уровнем кредитоспособности. В тоже время, премия за риск, в среднем, ниже. Однако для кредитоспособных заёмщиков процентная ставка кредита по-прежнему не зависит от индивидуальных характеристик (частной информации).

$$P_2(Q_i) = P_2, \text{ при } Q_i \in [Q_0, 1] \quad (1.2)$$

Банк № 3 классифицирует физических лиц по кредитоспособности по шкале с тремя градациями: “некредитоспособен”, “низкая кредитоспособность”, “высокая кредитоспособность”. Кредит выдаётся клиентам с “низкой кредитоспособностью” и “высокой кредитоспособностью”. Премия за риск для клиентов, принадлежащих к одному классу кредитоспособности, одинакова. Следовательно, существует две процентные ставки (1.3):

- процентная ставка для клиентов с классом “низкая кредитоспособность” ( $P'_3$ ) равна средней стоимости всех кредитов данного класса;
- процентная ставка для клиентов с классом “высокая кредитоспособность” ( $P''_3$ ) равна средней стоимости кредитов всех клиентов данного класса.

$$P_3(Q_i) = \begin{cases} P'_3, & \text{при } Q_i \in [Q_0, \bar{Q}) \\ P''_3, & \text{при } Q_i \in [\bar{Q}, 1] \end{cases} \quad (1.3)$$

В данном случае для заёмщиков процентная ставка по кредиту зависит от кредитоспособности, которая оценивается по частной информации. Однако количество клиентов банков достигает десятков миллионов, поэтому трёх градаций недостаточно, чтобы заёмщик ощутил влияние своего действительного уровня кредитоспособности на процентную ставку по кредиту, по этому подобная классификация является грубым инструментом воздействия на состав заёмщиков.

Применение банком шкалы с четырьмя градациями для классификации физических лиц по кредитоспособности позволит точнее дифференцировать процентную ставку по кредиту в соответствии с частной информацией клиентов. Пять градаций позволит банку ещё точнее дифференцировать процент-

ную ставку по кредиту. Увеличение числа градаций в пределе даёт непрерывную шкалу, что подробнее будет рассмотрено в следующей главе. Применение непрерывной шкалы позволяет банку определять индивидуальный класс кредитоспособности каждого клиента.

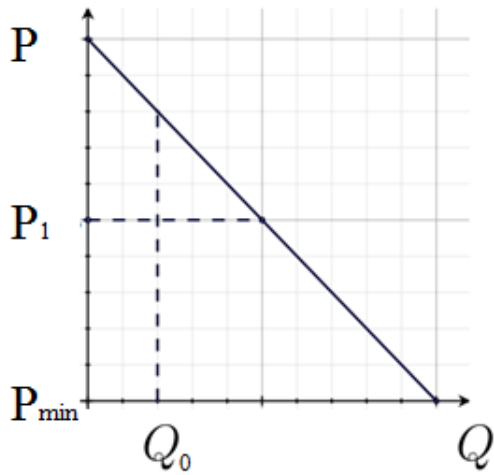
Банк № 4 классифицирует физических лиц по кредитоспособности по непрерывной шкале. Кредит выдаётся клиентам, у которых класс кредитоспособности выше минимально допустимого значения ( $Q_0$ ). Процентная ставка соответствует его стоимости кредита (1.4), то есть заёмщик с классом кредитоспособности  $Q_i$  получают процентную ставку  $P_4^i$ , с классом  $Q_j$  процентную ставку  $P_4^j$ , с классом  $Q_k$  процентную ставку  $P_4^k$  и т.д.

$$P_4(Q_i) = P(Q_i), \text{ при } Q_i \in [Q_0, 1] \quad (1.4)$$

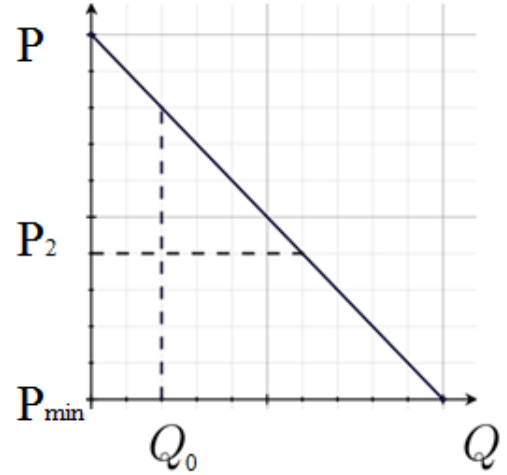
На Рисунках 1.4 - 1.9 показана зависимость процентной ставки от кредитоспособности клиентов. Уровень минимальной процентной ставки ( $P_{min}$  процентной ставки без кредитного риска) описан в третьей главе. По оси  $Q$  отмечена кредитоспособность клиентов. По оси  $P$  отмечены денежные единицы. Наклонная линия показывает зависимость себестоимости кредита от кредитоспособности. Согласно данной функции себестоимость кредита для заёмщика с идеальной кредитоспособностью ( $Q = 1$ ) равна минимальной процентной ставке ( $P_{min}$ ). Таким образом, значение процентной ставки по кредиту для клиентов в банках отличается в зависимости от применяемых методов классификация физических лиц по кредитоспособности и от действительного значения кредитоспособности.

Изначально заёмщики могут брать кредит только в банке, за которым они закреплены, то есть банки изолированы друг от друга. Изначальное состояние равновесия банков представлено на Рисунке 1.4.

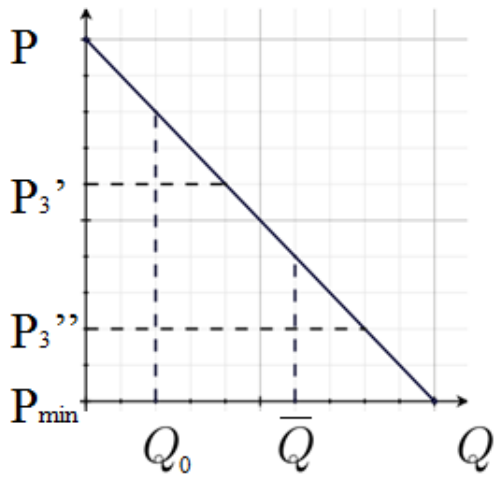
Рассмотрим попарную конкуренцию банков. Клиенты могут выбирать, в каком банке из рассматриваемой пары брать кредит. Клиенты обладают рациональным поведением, то есть они возьмут кредит в банке, который предоставит наименьшую процентную ставку.  $j$  - ый клиент возьмёт кредит в банке №  $n$ , если процентная ставка в данном банке ниже, чем в банке №  $m$  ( $P_n(Q_j) < P_m(Q_j)$ ). В противном случае данный клиент возьмёт кредит в бан-



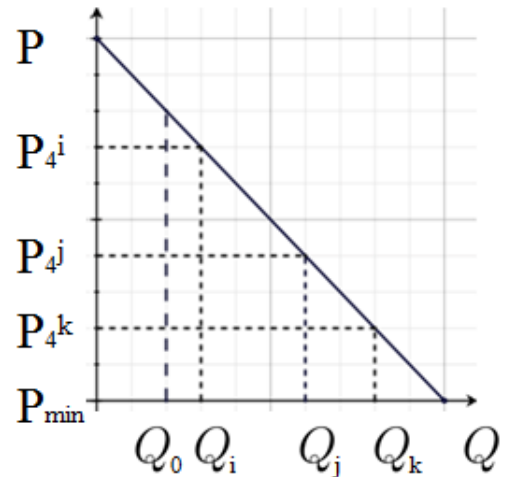
а) Банк № 1



б) Банк № 2



в) Банк № 3



г) Банк № 4

Рисунок 1.4: Изначальное состояние равновесия банков

ке №  $m$  ( $P_n(Q_j) > P_m(Q_j)$ ). Состояние равновесия достигнуто, когда ни один заёмщик не заинтересован перейти в другой банк ( $P_n(Q_j) = P_m(Q_j)$ ).

Допустим, банк № 1 и банк № 2 стали конкурировать на одном рынке. Заёмщики могут выбирать, в каком банке брать кредит. Процентная ставка в первом банке выше, чем во втором. Однако клиенты первого банка с классом кредитоспособности ниже минимально допустимого значения ( $Q_0$ ) удовлетворены своими условиями, потому что у них нет выбора. Банк № 2 для указанных клиентов закрыт, потому что для физических лиц с уровнем кредитоспособности ниже ( $Q_0$ ) кредит не предоставляют. Поэтому подобные клиенты второго банка перейдут в первый банк. Увеличение заёмщиков с низкой кредитоспособностью в первый банк приведёт к росту процентной ставки. В первом банке

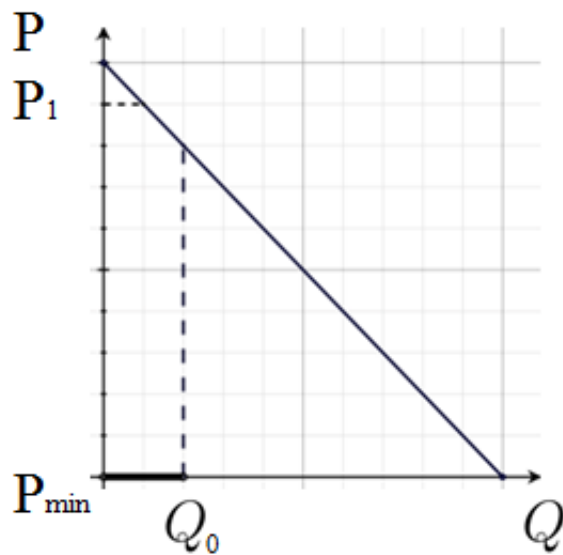
процентная ставка была изначально выше, чем во втором за счёт более низкого среднего уровня кредитоспособности. При перемещении клиентов из второго банка с недопустимым уровнем кредитоспособности, средний уровень кредитоспособности будет падать, а процентная ставка первого банка будет расти. Клиенты первого банка, в условиях, когда в первом банке процентная ставка изначально была выше, и тем более стала расти, начнут переходить во второй банк.

Возможно два варианта состояния равновесия:

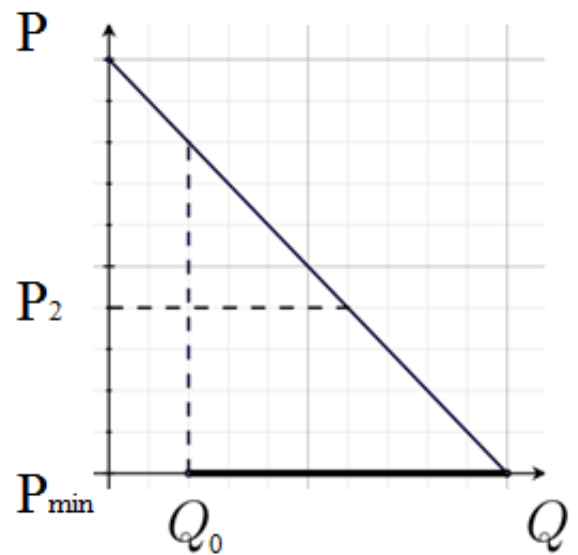
1. В банке № 1 окажутся все заёмщики с недопустимым уровнем кредитоспособности, а в банке № 2 все заёмщики с допустимым уровнем кредитоспособности (Рис. 1.5). В результате объединения в состоянии равновесия процентная ставка в банке № 1 выросла, а количество заёмщиков стало равным некредитоспособным клиентам из обоих банков. В банке № 2 процентная ставка осталась без изменения, зато увеличилось число заёмщиков до числа кредитоспособных клиентов из обоих банков.
2. В банке № 1 окажутся заёмщики преимущественно с очень высоким и очень низким уровнем риска, в том числе и клиенты с недопустимым во втором банке уровнем риска (Рис. 1.6). В банке № 2 будут находиться заёмщики с уровнем кредитоспособности, близким к среднему значению. Данная ситуация возможна в результате целенаправленного перехода заёмщиков из первого банка со значением кредитоспособности ниже среднего, но допустимым во втором банке. В случае подобного перехода процентная ставка в банке № 2 начнёт расти, а в банке № 1 падать.

Необходимо отметить, что для достижения второго варианта равновесия необходимо наличие достаточного числа клиентов с высоким уровнем кредитоспособности для того, чтобы процентные ставки в обоих банках смогли сравняться. В противном случае, процентная ставка в банке № 1 будет оставаться выше, чем в банке № 2, и постепенно все заёмщики с допустимым уровнем кредитоспособности перейдут в банк № 2.

Более того, второй вариант состояния равновесия, в отличие от первого, является неустойчивым к изменению объёма рынка. Допустим, произойдёт следующее из событий: заёмщик с кредитоспособностью выше среднего уровня уйдёт

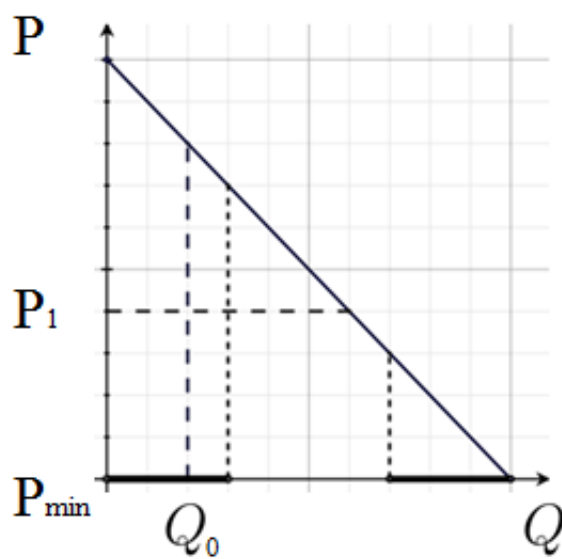


а) Банк № 1

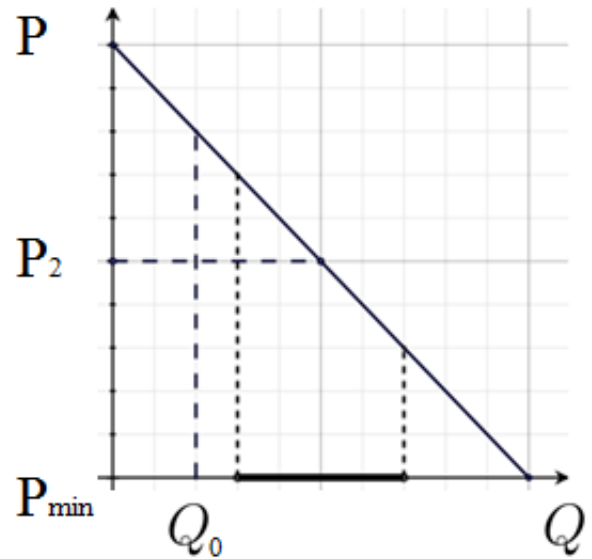


б) Банк № 2

Рисунок 1.5: Стабильное состояние равновесия при конкуренции между банком № 1 и банком № 2



а) Банк № 1



б) Банк № 2

Рисунок 1.6: Стабильное состояние равновесия при конкуренции между банком № 1 и банком № 2

из банка № 1; новый клиент с кредитоспособностью ниже среднего уровня придёт в банк № 1; заёмщик с кредитоспособностью ниже среднего уровня уйдёт из банка № 2; новый клиент с кредитоспособностью выше среднего уровня придёт в банк № 2.

При осуществлении любого из указанных событий, рынок при взаимодействии банка № 1 и № 2 выйдет из состояния равновесия. Для установления второго варианта равновесия в данном случае необходим переход из первого банка во второй заёмщика или заёмщиков с уровнем кредитоспособности ниже среднего значения, но выше допустимого в последнем. В противном случае рассматриваемый рынок перейдёт в первый вариант состояния равновесия. Из вышесказанного следует, что второй вариант равновесия достигается при особых условиях и является неустойчивым. В тоже время первый вариант равновесия является устойчивым. Если предположить, что сговор и целенаправленный переход клиентов отсутствует, то первый вариант равновесия будет наиболее вероятным.

Допустим, что при тех же начальных условиях, банк № 2 и банк № 3 стали конкурировать на одном рынке. Клиенты выбирают банк, который предложит наименьшую процентную ставку (1.5):

- Для заёмщиков с уровнем кредитоспособности от  $Q_0$  до  $\bar{Q}$  выгодно перейти из банка № 3 в банк № 2.
- Для заёмщиков с уровнем кредитоспособности выше  $\bar{Q}$  выгодно перейти из банка № 2 в банк № 3.

$$P_2(Q) \begin{cases} < P_3(Q), \text{ при } Q \in [Q_0, \bar{Q}] \\ > P_3(Q), \text{ при } Q \in [\bar{Q}, 1] \end{cases} \quad (1.5)$$

В процессе перехода клиентов средний уровень кредитоспособности в банке № 2 будет падать, а процентная ставка будет расти. В банке № 3 верхняя процентная ставка ( $P'_3$ ) и нижняя процентная ставка ( $P''_3$ ) будут снижаться, а средний уровень кредитоспособности ( $\bar{Q}$ ) будет перемещаться в сторону увеличения уровня кредитоспособности (Рис. 1.7).

Верхняя процентная ставка в банке № 3 окажется равна процентной ставке в банке № 2, в следующих случаях:

- когда клиенты с уровнем кредитоспособности выше минимально допустимого значения ( $Q_0$ ) и меньше среднего значения ( $\bar{Q}$ ) будут безразличны при выборе кредитора.

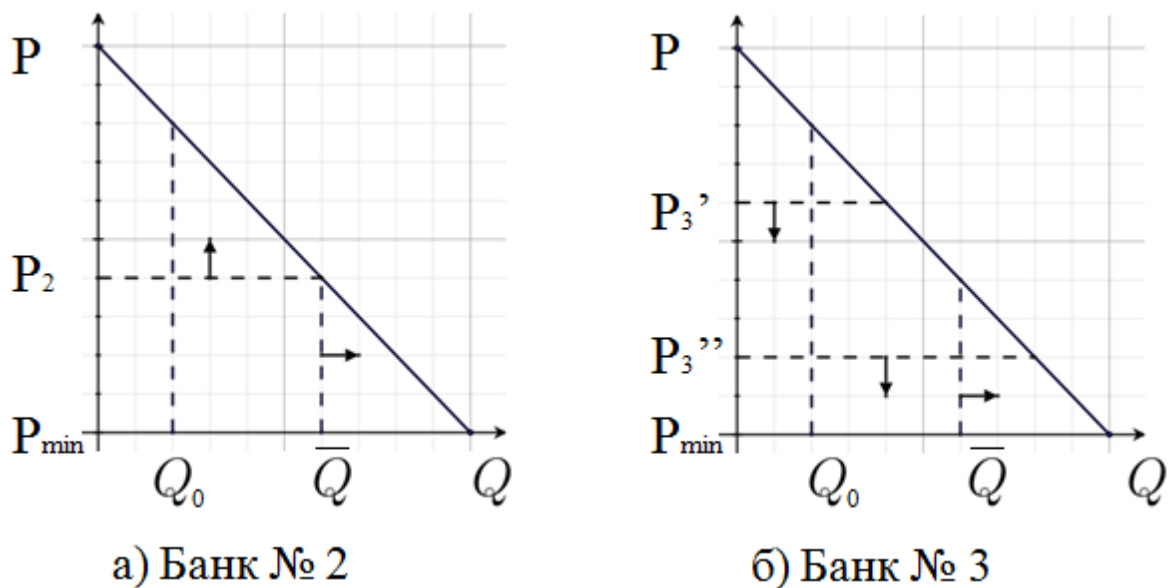


Рисунок 1.7: Динамика состава заёмщиков при конкуренции банка № 2 и банка № 3

- когда клиенты с уровнем кредитоспособности выше среднего значения ( $\bar{Q}$ ) будут заинтересованы брать кредит в банке № 3.

В случае, если в банке № 2 нет клиентов с уровнем кредитоспособности выше среднего уровня ( $\bar{Q}$ ), рассматриваемое состояние будет равновесным (Рис. 1.8).

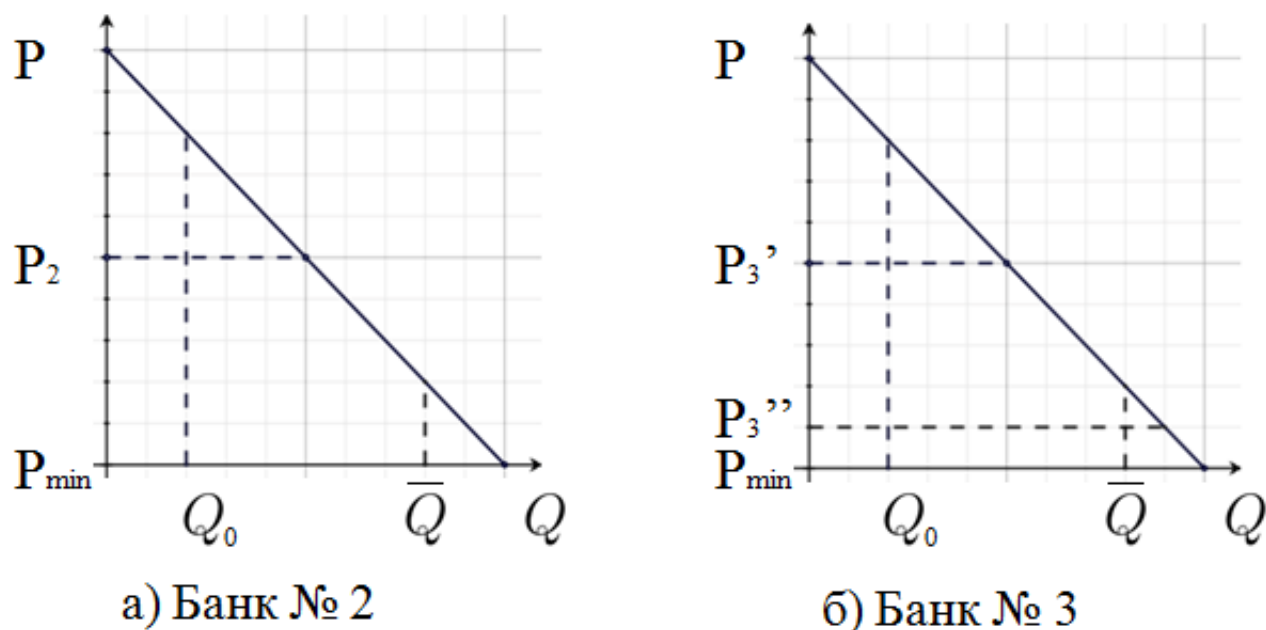


Рисунок 1.8: Нестабильное состояние равновесия при конкуренции банка № 2 и банка № 3

В данном равновесии средний уровень кредитоспособности заёмщиков во втором банке уменьшился, а процентная ставка увеличилась. В банке № 3,



напротив, данный уровень увеличился, а процентная ставка уменьшилась. В случае, если в данном состоянии равновесия в банке № 2 окажутся заёмщики с уровнем кредитоспособности выше среднего уровня ( $\bar{Q}$ ), то система выйдет из состояния равновесия. Получится, что процентная ставка в банке № 2 ( $P_2$ ) окажется выше верхней процентной ставки в банке № 3 ( $P'_3$ ). В таком случае, все заёмщики банка № 2 будут заинтересованы в переходе в банк № 1. Конечно, если из второго банка целенаправленно будут переходить в первый банк заёмщики с уровнем риска ниже среднего, то система сможет вернуться в состояние равновесия. Ранее было принято ограничение, что заёмщики не могут сговариваться, и переходят произвольно. Из данного ограничения следует, что система не сможет вернуться в рассматриваемое состояние равновесия. Более того, при данном ограничении рассмотренное ранее состояние не является равновесным, потому что, не сговариваясь, обязательно найдётся заёмщик в банке № 2 с уровнем кредитоспособности выше ( $\bar{Q}$ ), который выведет систему из состояния равновесия. В результате заёмщики будут переходить из банка № 2 в банк № 3. Процентная ставка в банке № 2 будет выше верхней процентной ставки в банке № 3. В состоянии равновесия в банке № 2 не окажется ни одного заёмщика.

При тех же начальных условиях рынок стал общим для третьего и четвертого банков. Как и в предыдущих вариантах объединения рынков, клиенты решают, в каком банке им брать кредит:

- при  $P_3(Q_i) < P_4(Q_i)$ , заёмщик  $i$  предпочтёт взять кредит в банке № 3;
- при  $P_3(Q_i) > P_4(Q_i)$ , заёмщик  $i$  предпочтёт взять кредит в банке № 4;
- при  $P_3(Q_i) = P_4(Q_i)$ , заёмщик  $i$  будет безразличен при выборе банка.

Множество клиентов, принадлежащих указанным выше группам, описано системой уравнений (1.6):

$$P_3(Q) \begin{cases} < P_4(Q), \text{ при } Q \in [Q_0, Q') \cup [\bar{Q}, Q'') \\ > P_4(Q), \text{ при } Q \in (Q', \bar{Q}) \cup (Q'', 1] \\ = P_4(Q), \text{ при } Q = Q' \cup Q = Q'' \end{cases} \quad (1.6)$$

Перемещение заёмщиков приведёт к тому, что процентная ставка в банке № 3 ( $P'_3$  и  $P''_3$ ) в среднем начнут движение вверх, а средний уровень кредитоспособности заёмщиков ( $\bar{Q}$ ) влево (Рис. 1.9). Чем выше процентные ставки в банке № 3, тем меньше заёмщиков заинтересовано брать в нём кредит. В результате в третьем банке окажутся 2 заёмщика с минимально допустимым уровнем риска по обоим банкам. Для данных заёмщиков средний уровень риска будет равным их индивидуальному уровню риска.

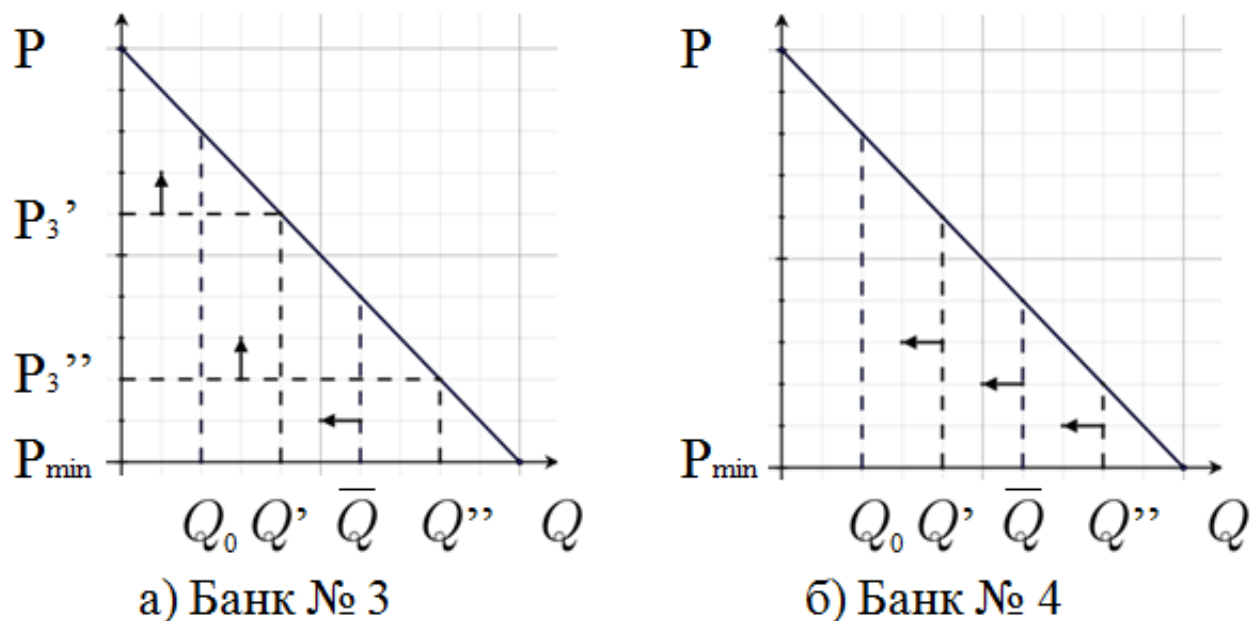


Рисунок 1.9: Динамика состава заёмщиков при конкуренции банка № 3 и банка № 4

В результате объединения в третьем банке осталось два заёмщика. Стоит отметить, что число оставшихся заёмщиков равно числу категорий процентных ставок. Предположим, что банк № 3 дифференцировал заёмщиков по уровню кредитоспособности, используя шкалу с четырьмя градациями. Группе заёмщиков с самым низким уровнем кредитоспособности кредит не выдаётся. Остальным трём группам выдается кредит. Процентная ставка усредняется по группе. Число градаций процентной ставки равно четырём. Банк № 3 и банк № 4 с некоторого момента стали доступны всем заёмщикам. Заёмщики выбирают банк с более низкой процентной ставки для них. Проделав выкладки, аналогичные вышеуказанным, рынок достигнет состояния равновесия, когда в третьем банке останутся четыре заёмщика с самой низкой оценкой кредитоспособности (самым высоким кредитным риском, начиная с допустимого уровня).

Было произведено сравнение банков, которые используют различные варианты дифференциации процентной ставки в соответствии с частной информацией клиента на рынке кредитования физических лиц. В примере рассмотрено 4 банка. Банки упорядочены по увеличению числа градаций при дифференциации клиентов банка:

- Банк № 1 не дифференцирует процентную ставку для клиентов.
- Банк № 2 использует частную информацию о заёмщиках для того, чтобы не выдавать кредиты для клиентов с наибольшим уровнем риска. Для заёмщиков процентная ставка не дифференцирована.
- Банк № 3 дифференцирует процентную ставку по шкале с двумя градациями.
- Банк № 4 дифференцирует процентную ставку по непрерывной шкале, то есть устанавливает индивидуальную процентную ставку для заёмщиков.

На основе конкуренции банков получены следующие результаты. Банк № 1 собирает всех заёмщиков с недопустимым уровнем кредитоспособности так как никто больше им кредит не выдаёт. Банк № 2 получает приток клиентов с допустимым уровнем кредитоспособности из банка № 1. Однако банк № 2 получает отток более надёжных заёмщиков (с более высокой кредитоспособностью) в пользу банка № 3. Банк № 4 при конкуренции с банком № 3 получает приток заёмщиков.

В состоянии равновесия в банке № 3 останется 2 заёмщика с наименьшим допустимым уровнем кредитоспособности. Допустим банк № 3 дифференцирует заёмщиков по шкале с четырьмя градациями, а процентную ставку по шкале с тремя. При конкуренции с банком № 4 в данном случае в банке № 3 останется три заёмщика с наименьшим допустимым уровнем кредитоспособности. Стоит отметить, что если в банке № 3 число заёмщиков совпадает с числом градаций процентной ставки или меньше, то при конкуренции с банком № 4 достигнуто состояние равновесия.

В результате попарной конкуренции банков со смежными номерами можно отметить, что клиенты с высокой кредитоспособностью стремятся взять кредит

в банке, где, при прочих равных условиях, больше градаций (классов), по которым дифференцирована процентная ставка. На первый взгляд, построенная модель конкуренции банков при различном отражении частной информации в процентная ставка может показаться идеализированной. Однако данная модель показывает, что не использование частной информации в процентной ставке даёт возможность конкуренту переманить клиентов за счёт более гибкого установления процентной ставки, в большей степени соответствующей ожидаемым затратам банка по кредиту. Результаты данной модели будут использованы при дальнейшей разработке ценовой стратегии, которая не будет приводить к неблагоприятному отбору заёмщиков на рынке потребительского кредитования.

### 1.3 Практика классификации физических лиц

В нынешнем параграфе рассматривается значимость оценки кредитного риска в потребительском кредитовании для экономики, а так же тенденция развития методов классификации физических лиц по кредитоспособности.

Банки анализируют информацию о клиентах для оценки кредитного риска. Кредитный риск – это возможность (potential) невыполнения заёмщиком своих обязательств [123]. Некоторые авторы вместо понятия возможность используют понятие правдоподобие (likelihood) [100], другие вероятность. Содержание состоит в том, что кредитный риск это некоторая мера, по которой банк ранжирует клиентов. Оценка кредитного риска основана на определении связи между характеристиками клиентов и вероятностью невыплаты кредита, полученная посредством анализа кредитных историй подобных клиентов банка [146]. Данное определение показывает, что оценка кредитного риска (классификация по кредитоспособности, кредитный скоринг) это одна из задач анализа данных (data mining), в основе которых лежит определение взаимосвязей в массиве данных [143]. Более того, кредитоспособность так же является мерой, по которой банк ранжирует клиентов. В книге Лин Томаса [135] кредитоспособность используется в качестве меры для определения кредитного риска клиента. Как было указано ранее кредитоспособность можно представить мерой, которая связана с вероятностью банкротства заёмщика некоторой убывающей функцией от кредитоспособности. В текущей работе кредитоспособность в широком смысле

– это способность клиента своевременно осуществлять выплаты по кредиту. Кредитоспособность в узком смысле – это мера, которую банк использует для вычисления процентной ставки по кредиту.

Основной деятельностью коммерческих банков является перераспределение временно свободных средств в экономике. Банки берут деньги на депозит и выдают кредиты. Существенное значение для функционирования банка имеет кредитный риск. Заёмщики имеют различный уровень кредитного риска, который не ограничен двумя значениями, как в большинстве методов классификации физических лиц по кредитному риску (кредитоспособности), что приводит к получению результата, обладающего низкой информативностью и к некачественному управлению кредитным риском.

В качестве одного из факторов мирового финансового кризиса 2008 г. принято считать провал рынка ипотечного кредитования. Особенностью ипотеки среди кредитов физическим лицам является большой размер выданных кредитных средств и длительный срок кредитования, что связано с крупными кредитными рисками. Чем выше кредитные риски, тем выше стоимость кредита при прочих равных условиях [64]. Для снижения финансовой нагрузки на население (процентной ставки) при взятии ипотеки в США существуют компании с государственным участием в ипотечном кредитовании, а так же используется секьюритизацию активов [137]. Секьюритизация создана для уменьшения риска банкротства и для получения более низкой процентной ставки от кредитора [82]. Секьюритизация ипотечного кредита при относительно низком спросе на недвижимость снижает процентную ставку [47]. Некоторые эксперты полагают, что секьюритизация ипотечного кредитования привела к финансовому кризису 2008 года в США [47]. При секьюритизации ипотечного кредита банки уделяют меньше внимания кредитоспособности потенциального заёмщика. Секьюритизация в ипотечном кредитовании уменьшает интерес банка в отборе и наблюдении за заёмщиками, что приводит к снижению качества портфеля ипотечных кредитов [119]. Желаемый эффект секьюритизации состоял в удешевлении ипотечного кредита. Вместо этого банки стали выдавать кредиты заёмщикам с более низким уровнем кредитоспособности, что в конечном итоге стало отправной точкой мирового финансового кризиса 2008 г. Поэтому учет кредитоспособности при выдаче кредита физическому лицу играет важ-

ную роль для стабильного развития экономики. Стоит отметить, что мировой финансовый кризис 2008 года весьма негативно отразился на российском производстве. Индекс промышленного производства за 2009 год снизился на 10,7%<sup>3</sup>.

В основе мирового финансового кризиса 2008 года некоторые эксперты видят не соответствующую действительности оценку кредитного риска, которая связана с некачественной работой кредитных агентств [63]. В 2010 году конгресс США принял акт Додда и Франка, который регулировал деятельность кредитных агентств с целью увеличения информативности кредитных рейтингов [63]. Однако результаты применения не оказали положительного эффекта. Объяснить данное явление можно тем, что кредитное агентство, как коммерческая организация, имеет цель максимизировать прибыль. После вступления в силу указанного выше акта кредитные агентства продолжили преследовать поставленную цель, однако к существующим затратам добавились расходы для удовлетворения новым требованиям, либо расходы на поиск путей по уклонению от удовлетворения данных требований. Аналогично в банковской практике существуют схемы обхода регулятивных требований [120].

Одним из современных направлений в экономической теории является теория инициатив [98]. Вместо добавления новых или ужесточения существующих требований кредитным агентствам, более эффективным представляется разработка системы взаимодействия участников на рынке потребительского кредитования, по которой:

- у банков или кредитных агентств будет инициатива (материальный стимул) предоставлять более информативную и точную оценку кредитного риска;
- у физических лиц будет инициатива предоставлять частную информацию в банк, что снизит затраты последнего на сбор информации.

В основе подобной системы могут быть модель и метод классификации физических лиц по кредитоспособности разрабатываемые в текущей работе. Разрабатываемые модель и метод классификации физических лиц по кредитоспособности, ориентированы на обеспечение возможности банку в автоматизированном режиме принимать больше управленческих решений для более каче-

---

<sup>3</sup>Электронный ресурс: [http://www.gks.ru/free\\_doc/new\\_site/business/prom/ind\\_prom\\_okved.xls](http://www.gks.ru/free_doc/new_site/business/prom/ind_prom_okved.xls)

ственного управления кредитным риском и стабилизацией рынка кредитования во избежание очередного мирового финансового кризиса.

Банки применяют финансовые показатели для отличия кредитоспособных заёмщиков от некредитоспособных, начиная с 1930-х годов [150]. Первой математической моделью оценки кредитоспособности физических лиц стала модель Дюрана [65]. Данная модель использовалась для принятия решения: выдавать кредит клиенту или отказать в выдаче кредита. По мере развития математического и инструментального аппарата, стали применяться иные методы для классификации физических лиц по кредитоспособности:

- метод линейного дискриминантного анализа (linear discriminant analysis) [118],
- метод логистической регрессии (logistic regression) [62, 132],
- метод опорных векторов (support vector machine) [43, 44, 56, 59, 79, 84, 138],
- метод нейронных сетей (neural networks) [86, 92, 141],
- метод  $k$  ближайших соседей (k-nearest neighbors) [81, 99],
- метод генетических алгоритмов (genetic algorithms) [52, 96],
- метод дерева решений (decision tree) [139],
- метод гауссовских процессов (gaussian process) [85],
- метод интеллектуальной иммунной системы [50]
- метод марковских цепей [49, 104].

Однако перечисленные методы по-прежнему позволяют оценивать кредитоспособность потенциальных заёмщиков по бинарной шкале: “кредитоспособен”, “не кредитоспособен”. Кроме того, методы могут быть предназначены для различных стадий классификации, таких как выбор факторов [111] или настройка параметров модели. В частности, выбор оптимального пограничного значения (cut off value) для отличия кредитоспособных клиентов от некредитоспособных. В итоге для осуществления требуемой классификации методы могут быть скомбинированы, например:

- методы нейронных сетей и генетических алгоритмов [116];
- методы нейронных сетей и опорных векторов [124];
- методы неточных множеств (rough sets) и дерева решений [149].

Комбинирование методов осуществляется для того, чтобы увеличить точность классификации потенциальных заёмщиков. Однако разработка более сложных методов классификации зачастую создает лишь видимость прогресса не увеличивая точности [77].

Существуют следующие определения кредитного скоринга (моделей и методов классификации физических лиц по кредитоспособности):

- статистическая модель оценки кредитоспособности потенциального заёмщика [78];
- метод измерения риска невозврата кредита клиентом на основе анализа его информации [79].

По производимому результату большинство методов классификации физических лиц по кредитоспособности схоже. Во многих работах кредитный скоринг представлен, как задача бинарной классификации [35, 37, 45, 58, 72, 78, 103, 113, 133]. Если банк анализирует кредитоспособность потенциального заёмщика по данным методам, то он получает один из двух результатов: кредитоспособен или некредитоспособен данный клиент.

Модель классификации физических лиц по кредитоспособности в работе [110] представлена, как отображение набора характеристик ( $X^d = (x_1, x_2, \dots, x_d)$ ) на бинарную шкалу (1.7).

$$f : X^d \rightarrow \{\text{кредитоспособен, некредитоспособен}\} \quad (1.7)$$

Учитывая, что некредитоспособным заёмщикам выдавать кредит не следует, как обладателям наибольшего кредитного риска, то все заёмщики, получающие кредит, относятся к одному классу кредитоспособных заёмщиков, что не позволяет банку на основе данного результата дифференцировать процентную ставку.



В последние годы менеджеры финансовых учреждений стали придавать существенное значение недостаточной точности существующих методов классификации физических лиц по кредитоспособности, что отчасти связано с возрастающим спросом на кредиты [140]. В тоже время в работах, посвященных разработке методов кредитного скоринга, авторы большое внимание уделяют точности классификации. Однако остаётся неясным, что авторы, разрабатывающие методы классификации физических лиц по кредитоспособности на основе бинарного подхода, подразумевают под точностью:

- достаточно ли точно относить в один класс клиентов, физических лиц, которые полностью не выплатили кредит с теми, кто один раз задержали платёж на 90 дней?
- корректно ли относить в один класс клиентов, физических лиц, которые ни разу не задержали платёж с теми, кто 5 раз задержал платёж менее 30 дней?

Бинарная классификация физических лиц по кредитоспособности обладает низкой информативностью для банка, то есть результат данной классификации не может быть использован для принятия обоснованных решений внутри класса кредитоспособных или некредитоспособных клиентов, соответственно. Некоторые авторы ставят вопрос: почему кредитные агентства используют небольшое число классов при классификации заёмщиков, в то время, как кредитный риск может принимать бесконечное множество значений [74]. Для построения меры, позволяющей отображать бесконечное множество необходимо использовать непрерывную шкалу. Таким образом, классификация клиентов по кредитоспособности должна осуществляться по непрерывной шкале, что позволит банку учесть любые изменения в кредитоспособности клиента.

Стоит отметить, что положение Банка России устанавливает пять классификационных категорий групп риска для заёмщиков и требует оценивать кредитный риск по выданным ссудам на постоянной основе [17]:

- стандартные;
- нестандартные;

- сомнительные;
- проблемные;
- безнадежные.

Каждая категория риска влечёт определённый уровень расходов, потому что по меньшей мере банкам необходимо формировать больше резервов под активы из более низкой категории качества. Банку по меньшей мере целесообразно дифференцировать процентную ставку в соответствии с данными расходами. Однако бинарной классификации недостаточно, чтобы в автоматизированном режиме установить подобное соответствие. Другими словами, учитывая, что некредитоспособным клиентам кредит не выдают, то одному классу кредитоспособных клиентов невозможно поставить в соответствие пять различных процентных ставок.

Стоит отметить, что с 01.01.14 к 01.07.15 увеличилась доля сомнительных кредитов с 6,9% до 8%, проблемных с 2% до 2,5% и безнадежных с 4% до 5,7% [1], что увеличивает актуальность проводимого исследования. В целом, применение банками именно автоматизированных систем классификации физических лиц по кредитоспособности обеспечивает существенный рост доходов банка [66]. В результате при использовании методов кредитного скоринга с бинарной классификацией потенциальных заёмщиков банк тратит средства на сбор информации о клиенте только для отсева наименее кредитоспособных, что не позволяет привлекать клиентов с высокой кредитоспособностью более низкими процентными ставками, а так же стимулировать последних предоставлять частную информацию в банк.

Стоит отметить, что ключевая функция финансовых организаций заключается в аккумулировании финансов и передаче их тем, чья доходность выше [129]. Для выполнения данной функции необходимо уметь оценивать доходность (в случае с кредитованием – кредитоспособность). Бинарная классификация физических лиц, используемая в большинстве банков, не способствует определению клиентов с высокой доходностью.

Оценка кредитоспособности юридических лиц содержит больше градаций нежели физических лиц. Данный факт связан с тем, что юридические лица

обязаны вести бухгалтерскую отчётность, по которой различные характеристики фирмы имеют стоимостную оценку, что упрощает анализ данных, а так же объем кредита, выдаваемый юридическим лицам, в среднем больше, что делает решения по выдаче кредита более дорогостоящим. Информацию физических лиц анализировать сложнее, поэтому вместо точной оценки банки устанавливают граничное значение допустимой кредитоспособности, не дифференцируя клиентов более точно.

Согласно закону о равных возможностях (Equal Credit Opportunity Act) 1974 года в США запрещено дискриминировать физических лиц при выдаче кредита по таким признакам, как национальность, раса и др. [121]. Однако по результатам работы [91] банки не удовлетворяют данному акту по расовому признаку. Подобная дискриминация является не конструктивной, потому что заявитель на кредит в данном случае ничего не может сделать, чтобы изменить свой кредитный статус. При наличии методов классификации физических лиц по кредитоспособности, учитывающей различные характеристики клиентов, на основе которой можно дифференцировать процентную ставку по кредиту, банки смогут не прибегать к грубой дискриминации заёмщиков, такой, как расовая.

Установление единой процентной ставки изначально было связано со сложностью классификации физических лиц по кредитоспособности более чем на два класса с достаточной точностью. Однако с учётом развития технологий, теперь помимо отсева некредитоспособных клиентов стоит задача установления процентной ставки в соответствии с результатом классификации по кредитоспособности. Точность классификации физических лиц по кредитоспособности определяет возможности банка использовать частную информацию для установления индивидуальной процентной ставки каждому заёмщику в соответствии с ожидаемыми потерями по ссуде. В результате от точности методов классификации физических лиц по кредитоспособности зависит качество управления кредитным риском.

Некоторые работы посвящены разработке моделей для периодической классификации фирм по кредитоспособности, чтобы информировать банк и заёмщиков об изменении финансового положения последнего [89, 126]. Учитывая, что число выданных кредитов в потребительском кредитовании выше при меньшем объёме, целесообразна разработка автоматизированной системы монито-

ринга за кредитоспособностью физических лиц с целью управления кредитным риском посредством корректировки условий кредитования, например по открытым кредитным линиям физическим лицам. Кредитный риск клиентов имеет бесконечное множество значений. Для того чтобы в автоматизированном режиме корректировать процентную ставку по кредиту в соответствии с кредитным риском (ожидаемыми потерями по ссуде) потенциального заёмщика необходимо использовать непрерывную функцию с непрерывным аргументом.

Вопрос определения оптимального количества градаций при построении оценки (классификации) объекта является предметом множества научных работ [55]. В последнее время появляются работы, в которых предлагаются модели классификации физических лиц по кредитоспособности отличающиеся от классической бинарной. В работе [94] авторы для увеличения информативности и практической ценности результата классификации предложили модель классификации физических лиц по кредитоспособности, где классы клиентов “хорошие” (кредитоспособные) и “плохие” (некредитоспособные) разбить на два класса каждый. В работе [53] предложена модель классификации физических лиц по кредитоспособности с использованием трех классов. Процентная ставка для заёмщиков по данной модели дифференцируется следующим образом:

- клиент с “низким кредитным риском” (высокая кредитоспособность) получает “дешёвый” кредит;
- клиент со “средним кредитным риском” (средняя кредитоспособность) получает “дорогой” кредит;
- клиент с высоким “кредитным риском” (низкая кредитоспособность, т.е. некредитоспособный) не получает кредит.

В работе [32] предложена модель классификации физических лиц по кредитоспособности по шкале с шестью классами. В работе [109] предложена модель классификации физических лиц по кредитоспособности по шкале с пятью классами:

1. лучшая (“best”) кредитоспособность,
2. хорошая (“good”) кредитоспособность,

3. удовлетворительная (“satisfactory”) кредитоспособность,
4. плохая (“bad”) кредитоспособность,
5. худшая (“worst”) кредитоспособность.

Одна из целей банка, как коммерческой организации, является максимизация прибыли собственников. Поэтому помимо качества управления кредитным риском, стоимость получения оценки кредитоспособности играет не мало-важную роль для коммерческих банков. Качество управления кредитным риском определяется точностью методов классификации физических лиц по кредитоспособности. Большинство подобных методов классификации физических лиц по кредитоспособности можно разделить на две группы, основанные на экспертном и статистическом подходе [144]. Преимущества автоматизированного кредитного скоринга по сравнению с субъективным принятием решений экспертами заключаются в следующем [110]:

- снижение затрат на процесс принятия решения по кредиту;
- производимый результат может быть отслежен и скорректирован;
- объективность полученных результатов.

Преимуществом экспертного подхода является высокая точность осуществляемого анализа кредитоспособности, достигаемая за счёт индивидуального оценивания, в котором учитываются множество нестандартных аспектов потенциального заёмщика. В данном случае, степень автоматизации при составлении оценки низкая, а требования к квалификации экспертов высокие. В результате использования экспертного подхода стоимость анализа кредитоспособности оказывается чрезвычайно велика. Кроме того, эксперты могут специализироваться на различных характеристиках клиента и включать такие специальности как:

- экономист для оценки финансового состояния клиента;
- психолог для оценки склонности клиента к мошенническим схемам;
- физиолог для оценки того, насколько клиент справляется с нагрузкой.

При использовании статистических методов степень автоматизации процесса высока за счёт стандартизации и усреднения характеристик потенциального заёмщика, что снижает качество результата, но во многих случаях позволяет заменить экспертов на операторов, в задачу которых входит обеспечение функционирования системы поддержки принятия решений. В результате стоимость анализа кредитоспособности при использовании статистических методов низкая, однако точность результата также оказывается ниже, нежели при экспертном подходе, что определяет более высокие затраты от некачественного управления кредитным риском.

При высокой стоимости анализа кредитоспособности, отрицательный эффект от увеличения организационных расходов банка в процентной ставке по кредиту может превысить положительный эффект от дифференциации заемщиков по уровню риска: снижения потерь по ссудам и повышения финансовой стабильности. Поэтому, наряду с методами классификации физических лиц по кредитоспособности, необходимы инструменты, позволяющие снизить стоимость анализа кредитоспособности.

Банки в среднем инвестируют 17% от суммы общих расходов на автоматизацию процессов [29]. Инструменты зависят от уровня развития технологий. Уровень технологий на данный момент находится в завершении пятого и в начале шестого технологического уклада [23]. В ядре пятого технологического уклада находятся вычислительная техника, программное обеспечение, робототехника, информационные технологии [4]. Иными словами, технологии, обеспечивающие автоматизацию процессов, что позволяет снизить затраты при выполнении ряда задач. В создаваемом шестом технологическом укладе автоматизация процессов происходит на более высоком уровне, что позволяет говорить о создании интеллектуальных информационных систем, которые позволяют выполнять задачи, свойственные творческой личности [29].

В условиях современного рынка необходимо инновационное развитие для обеспечения конкурентоспособности [16]. Приоритетным направлением при управлении кредитным риском является разработка системы, которая позволит классификации физических лиц по кредитоспособности с качеством сопоставимым с результатом группы экспертов, и в тоже время будет автоматизированной. Результат, полученный по данной системе, будет обладать высокой

точностью и низкой стоимостью. До тех пор пока оценка кредитоспособности производится по бинарной шкале, о высокой точности классификации физических лиц по кредитоспособности и качественном управлении кредитным риском говорить преждевременно. Поэтому в данной работе внимание сфокусировано в основном на шкале, по которой производится классификация физических лиц по кредитоспособности.

При разработке метода классификации физических лиц по кредитоспособности по непрерывной шкале целесообразно использовать методы теории нечётких множеств, потому что многие системы основанные на методах данной теории позволяют построить системы обладающие высокой точностью, близкой к экспертной и низкой стоимости за счет высокой степени автоматизации. Нечёткая логика (основа теории нечётких множеств) была представлена Л. Заде в 1965 г. [148]. В классической логике результат может быть представлен в бинарной форме {да, нет} или  $\{0, 1\}$ . В нечёткой логике используется весь диапазон от 0 до 1. Данный диапазон содержит континуум значений, как и кредитоспособность (кредитный риск) заёмщика. Нечёткая логика является обобщением классической логики, то есть любая система, заданная в классической логике, может быть задана в нечёткой логике без потери информации. Другой особенностью является то, что теория нечётких множеств позволяет использовать естественный язык (natural language) при разработке программ [147], что позволяет переводить опыт экспертов на машинный язык. Используя опыт высококвалифицированного эксперта построенная система может производить результат более качественный чем большинство экспертов.

Стоит отметить, что теория нечётких множеств приобретает широкое распространение в России и за границей, в частности даже при управлении производственными системами, где протекающие физические процессы имеют конкретные, чёткие значения [11]. При оценке кредитоспособности физических лиц применение нечёткой логики представляется более востребованным. Понятие кредитоспособность физического лица сложно формализовать и задать в терминах четкой логики, что не является препятствием в нечеткой логике при создании системы управления.

Нечеткая логика позволяет создавать системы, которые способны эффективнее многих других подходов сохранять информацию в нестандартных усло-

виях. Данное свойство связано с несколькими понятиями, одним из которых является нечеткое множество. Допустим, исследователь задал четкую модель “стола”: прямоугольная доска с четырьмя палками одинаковой длины. Анализируемым объектом является стол с закругленными углами. Автоматизированная система по классификации объектов, основанная на четкой модели “стола” определит данный объект, как “не стол”. Однако при использовании нечеткого множества “стол”, вывод будет примерно следующим: данный объект является “столом” со степенью принадлежности 0,95. Получается, что при использовании четкой модели была полностью утрачена информация о том, что данный объект, хоть и частично, но обладает свойством стола, а при нечеткой модели нет.

В результате для дальнейшей разработки системы поддержки принятия решений по выдаче кредита целесообразно использование теории нечетких множеств. Стоит отметить, что теория нечетких множеств применяется при анализе кредитоспособности физических лиц [25]. Однако большинство данных методов производят бинарную классификацию клиентов [108, 145], а некоторые – классификацию по шкале с тремя градациями [80] или пятью градациями [109, 144]. Усредняя результаты по дискретной шкале, данные методы в полной мере не используют преимущества нечеткой логики.

## 1.4 Обзор методов классификации объектов

Существует ряд часто используемых методов в кредитном скоринге таких как: логистическая регрессия, дискриминантный анализ, нейронные сети, метод опорных векторов. В рамках каждого из указанных методов существуют различные модификации. В текущем разделе представлены базовые версии рассматриваемых методов [45].

Часто используемым методом при классификации является регрессия, которая в общем виде является функцией от независимых переменных (1.8). Частным случаем является множественная линейная регрессия, где зависимая переменная является линейной комбинацией независимых (1.9).

$$y = F(x_1, x_2, \dots, x_n) \quad (1.8)$$



$$y = w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n \quad (1.9)$$

В кредитном скоринге одним из часто используемых является метод логистической регрессии. Задача состоит в следующем, обучающая выборка содержит  $N$  наблюдений  $\{(x_i, y_i)\}_{i=1}^N$ , состоящих из характеристик заёмщиков  $x_i \in R^n$ , набор которых соответствует одному из двух классов ( $y_i \in \{0,1\}$ ):

0 – некредитоспособный клиент;

1 – кредитоспособный клиент.

Далее с помощью логистической функции (1.10) определяется вероятность того, что потенциальный заёмщик с набором характеристик  $x = \{x_1, x_2, \dots, x_m\}$  соответствует классу кредитоспособных клиентов (Рис. 1.10).

$$P(y = 1|x) = \frac{1}{1 + e^{-(w_0 + w^T x)}} \quad (1.10)$$

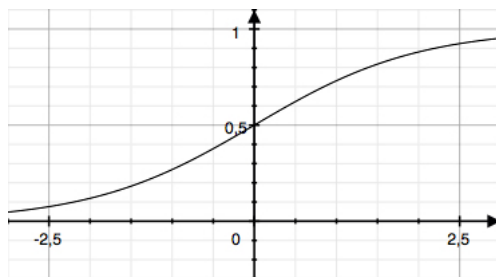


Рисунок 1.10: Функция логистической регрессии

Коэффициенты регрессии определяются по методу максимального правдоподобия. Для этого составляется функция правдоподобия (1.11), которая выражает плотность вероятности совместного осуществления событий  $Y_1, Y_2, \dots, Y_m$ .

$$L(Y_1, Y_2, \dots, Y_m; \theta) = p(Y_1; \theta) \cdot p(Y_2; \theta) \cdot \dots \cdot p(Y_m; \theta) \quad (1.11)$$

Далее необходимо найти такое значение  $\Theta$  (1.12), которое максимизирует функцию правдоподобия. На практике более эффективным является нахождение максимума логарифма функции правдоподобия (1.13), учитывая, что максимум обе функции принимают при одинаковых значениях аргументов.

$$\Theta = \Theta(Y_1, Y_2, \dots, Y_m) \quad (1.12)$$

$$\ln(L(Y; \theta)) \rightarrow \max \quad (1.13)$$

Другим распространённым в кредитном скоринге является метод дискриминантного анализа, по которому вычисляется апостериорная вероятность  $p(y|x)$  того, что наблюдение  $x$  принадлежит одному из классов  $y \in \{0,1\}$ . Вычисление постериорной вероятности осуществляется с помощью теоремы Байеса (1.14).

$$p(y|x) = \frac{p(y|x)p(y)}{p(x)} \quad (1.14)$$

В случае, когда условное распределение является Гауссовским (1.15), где  $\mu_1$  вектор средних значений класса 1, а  $\Sigma_1$  ковариационная матрица класса 1. Классификационное правило заключается в том, что если выполняется условие (1.16) тогда значения  $x$  соответствуют классу 1, в противном случае 0.

$$p(x|y = 1) = \frac{1}{(2\pi)^{n/2}|\Sigma_1|^{1/2}} \times e^{-\frac{1}{2}(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1)} \quad (1.15)$$

$$\begin{aligned} (x - \mu_1)^T \Sigma_1^{-1} (x - \mu_1) - (x - \mu_0)^T \Sigma_0^{-1} (x - \mu_0) < \\ < 2(\log(P(y = 1)) - \log(P(y = 0))) + \log|\Sigma_0| - \log|\Sigma_1| \end{aligned} \quad (1.16)$$

В связи с наличием квадратичных форм  $(x^T \Sigma_1^{-1} x, x^T \Sigma_0^{-1} x)$  метод называется квадратичным дискриминантным анализом (quadratic discriminant analysis). В случае, если  $\Sigma_0 = \Sigma_1 = \Sigma$  квадратичные формы исчезают и классификационное правило становится линейным по  $x$  и соответствующий метод называется линейным дискриминантным анализом (linear discriminant analysis).

Другим методом, используемым в кредитном скоринге, является линейное программирование, которое формулируется, как задача минимизации суммы ошибок (1.17) при ограничениях (1.18).

$$\min_{w, \xi} \sum_{i=1}^N \xi_i \quad (1.17)$$

$$\begin{cases} y_i[w^T x_i] \geq c - \xi_i, & y_i = +1 \\ y_i[w^T x_i] \leq c + \xi_i, & y_i = -1 \\ \xi_i \geq 0, & i = 1, \dots, N \end{cases} \quad (1.18)$$

Система неравенств предназначена для того, чтобы разделить кредитоспособных от некредитоспособных клиентов на заданном уровне  $c$ . Образно говоря, в методе линейного программирования определяется гиперплоскость, отделяющая классы наблюдений. Проблема заключается в том, что во многих случаях данных линий может быть проведено бесконечное множество (Рис. 1.11). Решение данной проблемы представлено в методе опорных векторов.

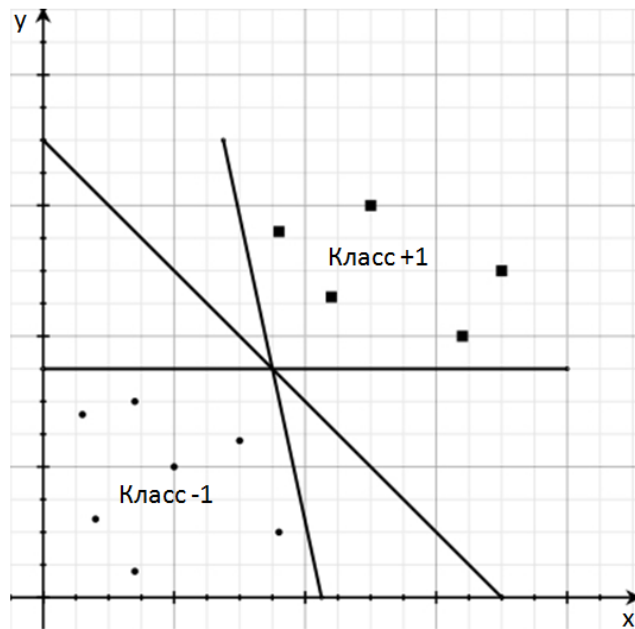


Рисунок 1.11: Проведение различных разделителей множеств по методу линейного программирования

В методе опорных векторов обучающая выборка состоит из  $N$  наблюдений  $\{(x_i, y_i)\}_{i=1}^N$ , состоящих из исходных наблюдений  $x_i \in R^n$  двух классов  $y \in \{-1, +1\}$ . Условие формулируется в виде системы неравенств (1.19) (можно записать в более кратком виде (1.20))

$$\begin{cases} w^T \varphi(x_i) + b \geq +1, & y_i = +1 \\ w^T \varphi(x_i) + b \leq -1, & y_i = -1 \end{cases} \quad (1.19)$$

$$y_i[w^T \varphi(x_i) + b] \geq 1, \quad i = 1, \dots, N \quad (1.20)$$

Нелинейная функция  $\varphi(\cdot)$  переводит исходные значения в пространство признаков. В данном пространстве признаков система неравенств (1.19) определяет гиперплоскость (1.21), которая разделяет классы. Далее оптимизируется функция (1.22) при ограничениях (1.23). Другими словами определяются параллельные гиперплоскости с максимальным расстоянием между классами.

$$w^T \varphi(x_i) + b = 0 \quad (1.21)$$

$$\min_{w,b,\xi} \tau(w,b,\xi) = \frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i \quad (1.22)$$

$$\begin{cases} y_i [w^T \varphi(x_i) + b] \geq 1 - \xi_i, & i = 1, \dots, N \\ \xi_i \geq 0, & i = 1, \dots, N \end{cases} \quad (1.23)$$

Переменные  $\xi_i$  являются дополнительными и необходимы для того, чтобы допустить возможность неверной классификации признаков по системе неравенств (например из-за пересекающихся распределений). Содержательно превосходство метода опорных векторов от линейного программирования заключается в отделении классов двумя параллельными гиперплоскостями, расстояние между которыми максимально (1.24), где  $w^T \varphi(x) + b = 0$  это гиперплоскость, равноудалённая от каждого из классов.

$$y(x) = \text{sign}(w^T \varphi(x) + b) \quad (1.24)$$

Для решения поставленной оптимизационной задачи (1.22, 1.23) составляется функция Лагранжа (1.25). Для построенной функции Лагранжа необходимо найти седловую точку, минимизируя по  $w, b, \xi$  и максимизируя по  $\alpha$  и  $v$  (1.26).

$$\Lambda(w, b, \xi; \alpha, v) = \tau(w, b, \xi) - \sum_{i=1}^N \alpha_i (w^T \varphi(x_i) + b - 1 + \xi_i) - \sum_{i=1}^N v_i \xi_i \quad (1.25)$$

$$\max_{\alpha, v} \min_{w, b, \xi} \Lambda(w, b, \xi; \alpha, v) \quad (1.26)$$

Результат получен после решения системы дифференциальных уравнений (1.27). Подставив решение для  $w$  в (1.24), получится значение для классификатора (1.28).

$$\begin{cases} \frac{\partial \Lambda}{\partial w} = 0 \longrightarrow w = \sum_{i=1}^N \alpha_i y_i \varphi(x_i) \\ \frac{\partial \Lambda}{\partial b} = 0 \longrightarrow \sum_{i=1}^N \alpha_i y_i = 0 \\ \frac{\partial \Lambda}{\partial \xi} = 0 \longrightarrow 0 \leq \alpha_i \leq C, i = 1, \dots, N \end{cases} \quad (1.27)$$

$$y(x) = \text{sign}\left(\sum_{i=1}^N \alpha_i y_i K(x_i, x) + b\right) \quad (1.28)$$

Графически содержание метода опорных векторов заключается в том, чтобы максимизировать расстояние между гиперплоскостями, отделяющими разные классы (Рис. 1.12). В этом состоит одно из отличий от линейного программирования, где гиперплоскость, разделяющую классы наблюдений, можно было построить под разными углами.

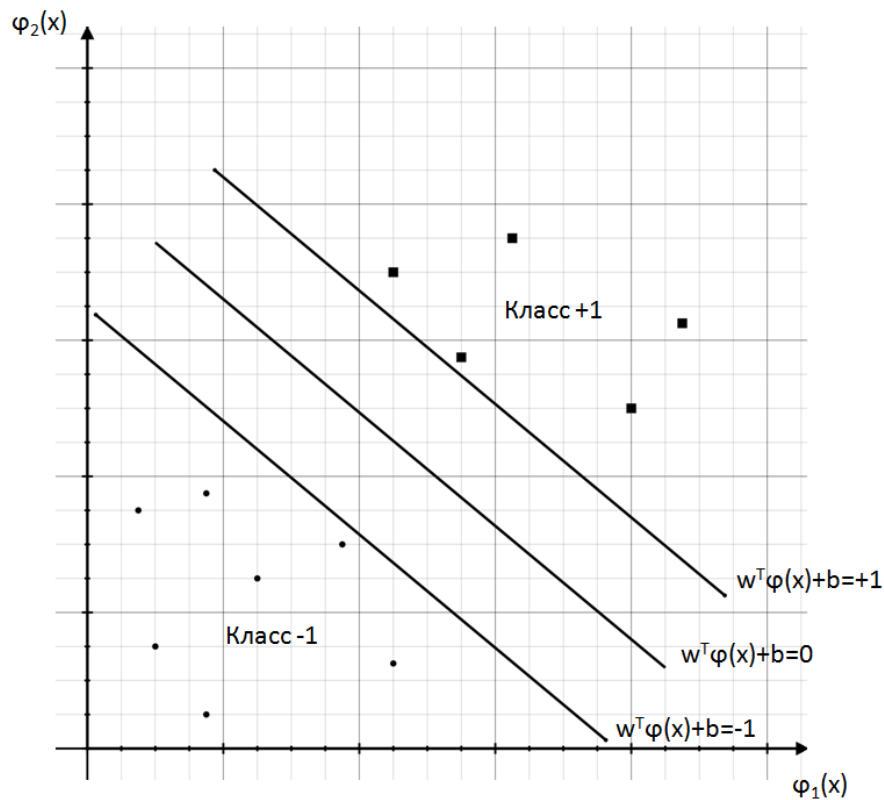


Рисунок 1.12: Классификация объектов по методу опорных векторов

В кредитном скоринге также применяется метод  $K$  ближайших соседей, который позволяет классифицировать наблюдение по  $K$  наиболее близким наблю-

дениям рассматриваемого множества. Для исключения зацикленности алгоритма достаточно  $K$  выбирать нечётным. Наиболее часто используется Эвклидова мера для определения ближайших значений (1.29), где  $x_i, x_j \in R^n$  значения наблюдений  $i$  и  $j$  соответственно.

$$d(x_i, x_j) = \|x_i - x_j\| = [(x_i - x_j)^T(x_i - x_j)]^{1/2} \quad (1.29)$$

При кластеризации по методу  $K$  ближайших соседей последовательность действий графически выглядит следующим образом. Сначала исходные  $m$  наблюдений объединяются в  $l$  кластеров ( $l < m$ ), определяемых расстоянием друг от друга. Вычисляются координаты кластера, например по методу центра тяжести (1.30) [61]:

$$\bar{d}_i(k) = \frac{\sum_{j=1}^k x_{ij}}{k}, \quad (1.30)$$

где,  $\bar{d}_i(k)$  – центр  $i$ -го кластера, содержащего  $k$  элементов, а  $x_{ij}$  –  $j$ -ый элемент  $i$ -го кластера.

Полученные кластеры объединяются в  $k$  кластеров ( $k < l$ ) и так далее до достижения необходимого уровня кластеризации. В случае с бинарной оценкой кредитоспособности конечное число кластеров равно 2. Кластер нового наблюдения (нового клиента, подавшего заявку на кредит) определяется по схожести характеристик с имеющимися кластерами. Например, для определения принадлежности нового наблюдения конкретному кластеру можно использовать меру  $d$  (1.31) [30]. Стоит отметить, что статистические методы являются весьма чувствительными к выбросам в исходных данных [112].

$$d(T_i, S_i) = \sqrt{\sum_{i=1}^F w_i (T_i - S_i)^2} \quad (1.31)$$

Другим используемым методом в кредитном скоринге является метод нейронных сетей, который является математическим представлением взаимодействия нейронов в мозге. Существует множество разновидностей метода нейронных сетей. Наиболее часто используемая разновидность при решении задач классификации является многослойная нейронная сеть (Multilayer Perceptron).

По структуре многослойная нейронная сеть состоит из входного слоя, выходного слоя и одного или нескольких скрытых слоёв. Скрытый слой состоит из нейронов, каждый из которых имеет по меньшей мере один вход из предыдущего слоя и генерирует по меньшей мере один выход в следующий слой сети (Рис. 1.13). Результат, производимый нейроном, вычисляется посредством взвешивания входных величин с учётом смещения  $b_i^{(1)}$  (1.32).

$$h_i = f^{(1)}(b_i^{(1)} + \sum_{j=1}^n W_{ij}x_j) \quad (1.32)$$

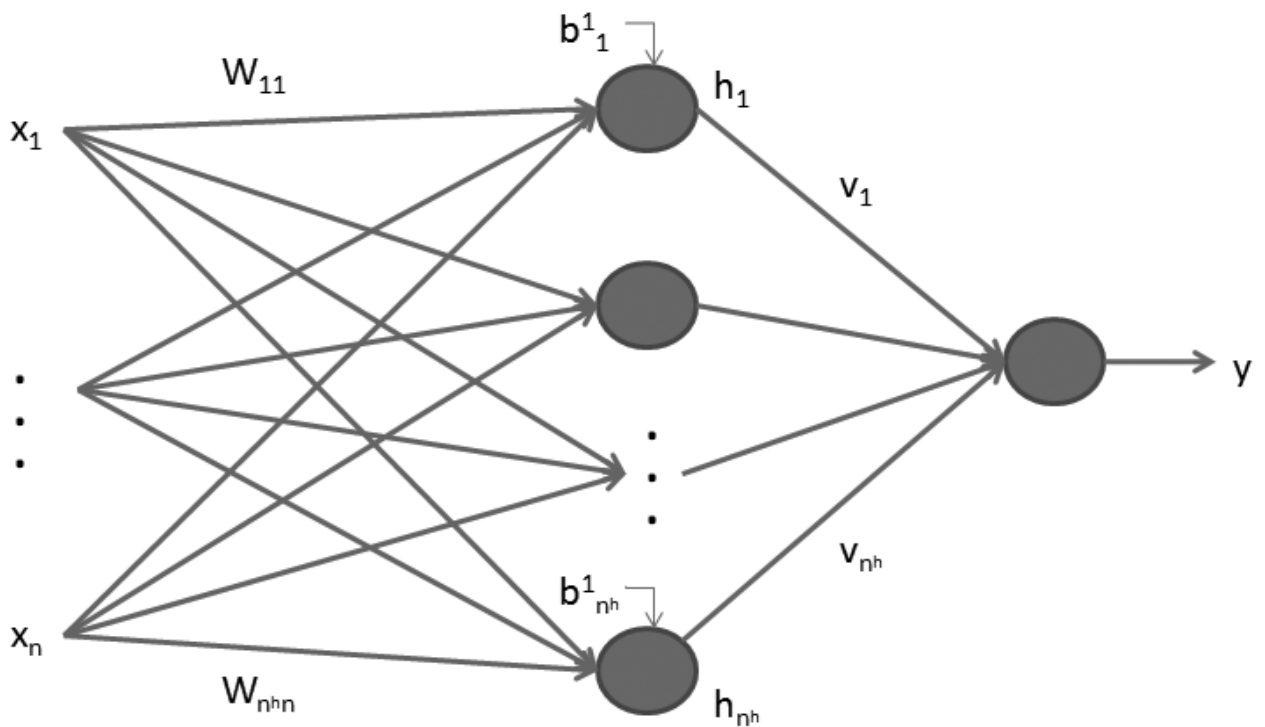


Рисунок 1.13: Нейронная сеть

Результат выходного слоя вычисляется по формуле (1.33):

$$y = f^{(2)}(b^{(2)} + \sum_{j=1}^{n_h} v_j h_j) \quad (1.33)$$

где  $n_h$  – количество скрытых нейронов,  $v$  – вектор весовых коэффициентов.

В качестве преобразующей функции ( $f$ ) в методе нейронных сетей часто используют сигмоидальную (1.34), гиперболический тангенс (1.35) и линейную функции (1.36). Стоит отметить, что в методе нейронных сетей существует

сложность в интерпретации результатов [70], а так же при маленьких выборках несущественные параметры могут иметь большой вес в итоговой оценке [38].

$$f(x) = \frac{1}{1 + e^{-x}} \quad (1.34)$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (1.35)$$

$$f(x) = x \quad (1.36)$$

Также одним из методов, используемых в кредитном скоринге, является дерево решений. Содержание данного метода заключается в представлении альтернатив в виде дерева (Рис. 1.14).

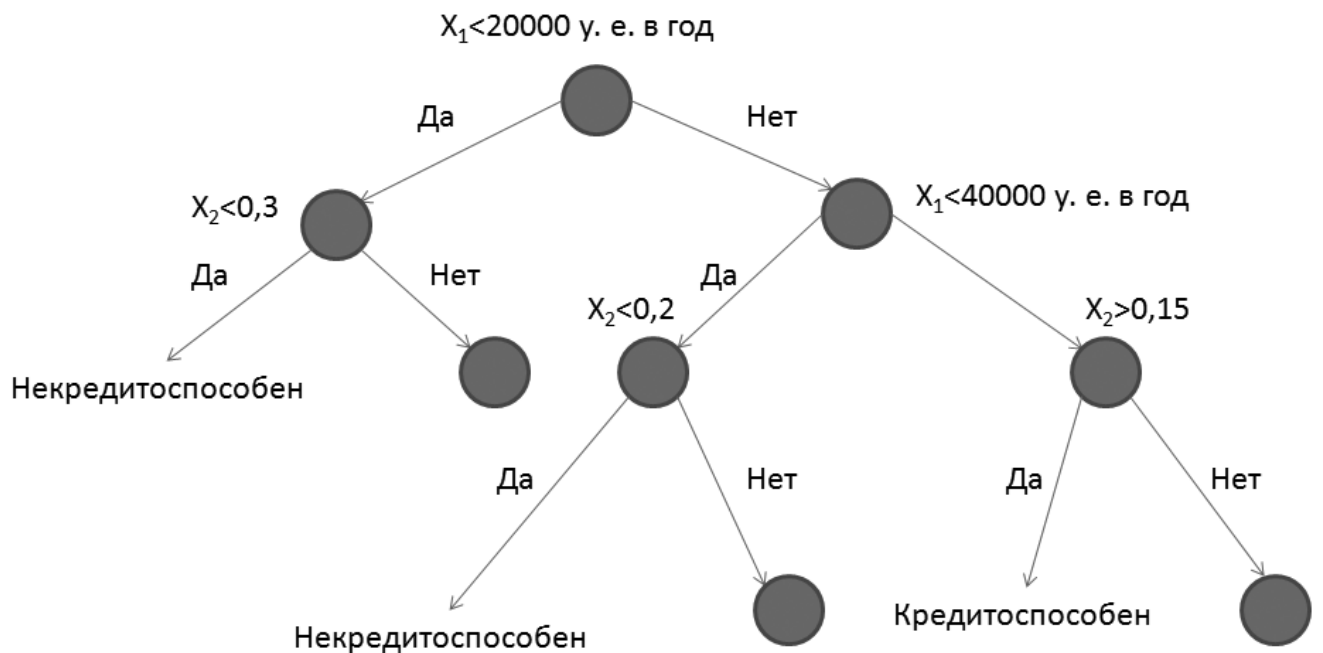


Рисунок 1.14: Пример дерева решений для кредитного скоринга

В кредитном скоринге так же используются и экспертные методы, например метод анализа иерархий [13]. В указанной работе данный метод используется как основа при вычислении дальнейших параметров модели. Согласно методу анализа иерархий, лицо, принимающее решение, вместо непосредственного построения искомой оценки занимается оценкой составных элементов с последующим объединением результатов (1.37) [14].



$$U(y) = \sum_{i=1}^m w_i y_i \quad (1.37)$$

Для начала следует подробно объяснить в чем заключается иерархия. Например, задача состоит в выборе депутата госдумы. Есть определённый набор критериев, которым депутат должен удовлетворять, а так же набор кандидатов, что в совокупности может быть записано в виде трёхуровневой иерархии.

Среди иерархий выделяют полную и неполную. Полная иерархия означает, что между всеми элементами смежных уровней иерархии существует связь. Таким образом, если взять элемент и все связанные с ним элементы более низкого уровня иерархии, то опять получится полная иерархия. Возвращаясь к примеру с выбором депутата. Допустим, что избирается председатель госдумы. Данная задача может быть записана в виде четырёхуровневой иерархии. Причем, если критерии выбора председателя и всех депутатов одинаковы, то полученная четырёхуровневая иерархия является полной. Модель может стать и пятиуровневой, если для председателя применяется новый набор критериев.

Неполная иерархия означает, что существует по меньшей мере одна пара элементов смежных уровней, для которых отсутствует связь. Стоит отметить, что упомянутую ранее задачу выбора председателя можно записать в виде неполной иерархии. Например, неполной иерархия будет, если каждый депутат избирается только в одном округе.

По методу анализа иерархий можно определить значения весовых коэффициентов (1.38). При применении метода анализа иерархий подразумевается, что эксперту сложно задать точное значение весовых коэффициентов, но он в состоянии попарно сравнивать все характеристики по 17-ти бальной шкале (1.39). Эксперту вместо  $m$  вопросов о значимости коэффициентов приходится отвечать на  $m * (m - 1) / 2$  вопросов о соотношении характеристик, определяя величины  $a_{ij}$  для всех пар характеристик.

$$w = (w_1, w_2, \dots, w_m) \quad (1.38)$$

$$a_{ij} \in \left\{ \frac{1}{9}, \frac{1}{8}, \dots, \frac{1}{2}, 1, 2, \dots, 8, 9 \right\}, a_{ij} = 1/a_{ji}, i, j = 1, 2, \dots, m \quad (1.39)$$

Информация, данная экспертом, представлена в виде матрицы попарных сравнений (1.40). Данная матрица содержит избыточную информацию, которая используется для определения логичности эксперта. Абсолютно логичный эксперт представляет информацию, соответствующую величине (1.41).

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix} \quad (1.40)$$

$$a_{ij} = \frac{w_i}{w_j}, i, j = 1, 2, \dots, m \quad (1.41)$$

Обозначим матрицу, составленную абсолютно логичным экспертом  $A^0$ . В таком случае выполняется соотношение (1.42), в котором  $w$  - является собственным вектором матрицы  $A^0$ , а  $\lambda_1$  - собственным числом. Строки матрицы  $A^0$  пропорциональны, то есть ранг матрицы равен единице. Величину  $\delta$  (1.43) можно использовать для оценки логичности ответов эксперта об относительной значимости характеристик. При малых значениях отклонения полученные результаты считаются удовлетворительными, а собственный вектор  $w$  используют в качестве весовых коэффициентов.

$$A^0 w = m w \quad (1.42)$$

$$\delta = \frac{\lambda_{max} - m}{m - 1} \quad (1.43)$$

Стоит отметить, что нелинейные методы зачастую показывают лучший результат классификации по сравнению с линейными. Однако во многих подобных методах затраты времени на обработку данных экспоненциально зависят от объема выборки [79].

В литературе по кредитному скорингу (методам классификации объектов, применяемых для определения кредитоспособности физических лиц) не существует единого мнения по поводу того, какой метод позволяет получать более качественную оценку кредитоспособности. Исследователи получают высокие результаты эффективности собственных методов. Однако данное явление часто

связано с тем, что исследователь является специалистом в разрабатываемом методе, а так же уделяет больше внимания настройке параметров данного метода, нежели сравниваемого [134]. Здесь проявляется конфликт интересов: исследователь заинтересован в том, чтобы показать высокое качество разрабатываемого метода и может даже неосознанно искажать результаты исследования в пользу собственного метода. Разрабатываются новые, всё более сложные методы кредитного скоринга, которые по факту создают лишь видимость прогресса, потому что зачастую простые методы позволяют получить более качественный результат [77]. В статье [33] проведён анализ 214 работ по кредитному скорингу (статей, книг и докладов), на основании которых авторы пришли к выводу, что не существует идеального метода кредитного скоринга.

Из вышеперечисленных методов только метод анализа иерархий позволяет построить оценку по непрерывной шкале. Однако метод анализа иерархий налагает жёсткие ограничения на информацию, предоставляемую экспертами. Каждая характеристика должна пройти процедуру попарного сравнения с остальными, в результате которого каждой паре ставится оценка по 17-ти бальной шкале. Больше возможностей в анализе дополнительной информации об исходных характеристиках представляет метод рандомизированных сводных показателей. Как и в методе анализа иерархий, метод рандомизированных показателей может использовать информацию о попарном доминировании характеристик. Однако метод рандомизированных сводных показателей позволяет строить оценку и при отсутствии информации о попарном доминировании, используя как экспертную, так и другие источники информации о значимости характеристик. Метод рандомизированных сводных показателей имеет широкий диапазон применимости для построения оценок [40]. На основе метода рандомизированных сводных показателей будет построена оценка кредитоспособности физических лиц по непрерывной шкале. Стоит отметить, что экспертные системы показывают высокие результаты при анализе финансового сектора [41], что указывает на перспективность применения метода рандомизированных сводных показателей в секторе кредитования физических лиц.

## Результаты главы

В данной главе была представлена текущая характеристика рынка потребительского кредитования России. Рассматриваемый рынок продемонстрировал высокую нестабильность при колебаниях рыночной конъюнктуры за 2014 и 2015 гг. Установление единой процентной ставки для всех кредитоспособных физических лиц не способствует стабилизации рынка. Банкам необходимо дифференцировать процентную ставку в соответствии с кредитоспособностью клиентов, что требует наличие подходящих моделей и методов.

Проведенный обзор моделей и методов классификации физических лиц по кредитоспособности показал, что большинство подобных моделей и методов предназначены для классификации физических лиц по бинарной шкале, что не позволяет в автоматизированном режиме на основе результата классификации дифференцировать процентную ставку по кредиту. Кредитоспособность клиентов может принимать бесконечное множество значений, что может быть учтено в процентной ставке с помощью классификации клиентов по кредитоспособности по непрерывной шкале. Таким образом, разработка и применение модели и метода классификации физических лиц по степени кредитоспособности по непрерывной шкале необходима для обеспечения конкурентного преимущества банку для привлечения клиентов с высокой кредитоспособностью, потому что указанные клиенты смогут получить более низкую процентную ставку, а так же для стабилизации рынка кредитования физических лиц.

Более того, в основе мирового финансового кризиса 2008 г. было проявление безразличия банков к кредитоспособности клиентов при принятии решения о выдаче кредита в секторе ипотечного кредитования. Финансовый кризис негативно сказался на реальном секторе экономики России. Таким образом, учёт кредитоспособности клиента при выдаче кредита необходим не только для обеспечения стабильности рынка кредитования, но и для всей экономики.

## Глава 2.

# Разработка метода классификации объектов по непрерывной шкале

В предыдущей главе была обоснована необходимость классификации физических лиц по кредитоспособности по непрерывной шкале. Был проведён обзор методов классификации объектов. В текущей главе построены модель и метод для классификации объектов по непрерывной шкале, а так же приведён пример осуществления искомой классификации при кредитовании физических лиц.

## 2.1 Модель классификации объектов по непрерывной шкале

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.

**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  $W$ .

Приведенную модель можно использовать для анализа кредитоспособности физических лиц, составив векторы характеристик объектов из частной информации последних (возраст, доход и т.д.). Искомый показатель кредитоспособности  $Q(x, w)$  необходимо строить на основе анализа потерь по ссудам физическим лицам. Классификация физических по кредитоспособности по непрерывной шкале может быть использовано банком для решения следующих задач:

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

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 ( $c$ ) (2.1).

$$P = P_{min} + c \quad (2.1)$$

Способ выражения премии за кредитный риск через класс кредитоспособности представлен в описании положения 4 основных научных результатов (2.2).

$$c(Q) = \sum_{i=0}^k a_i Q^i, \quad a_i \in R, \forall i = 0, 1, 2, \dots, k \quad (2.2)$$

Таким образом процентную ставку по кредиту для  $j$ -го клиента можно определить через класс кредитоспособности (2.3).

$$P(Q_j) = P_{min} + \sum_{i=0}^k a_i Q_j^i \quad (2.3)$$

Премия за кредитный риск может принимать значения в следующем 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:  $P(Q) \in [P_{min}, P_{min} + c_{max}]$ , в отличие от существующих моделей.

Наиболее распространённая модель классификации физических лиц по кредитоспособности является бинарной (2.4). 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.

$$Q^{(2)}(x, w) \in \{0, 1\} \quad (2.4)$$

Существуют модели, в которых классификация физических лиц по кредитоспособности осуществляется с использованием трех (2.5) и пяти классов (2.6), соответственно.

$$Q^{(3)}(x, w) \in \{0, \frac{1}{2}, 1\} \quad (2.5)$$

$$Q^{(5)}(x, w) \in \{0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, 1\} \quad (2.6)$$

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 (2.7).

$$Q^{(n)}(x, w) \in \{0, \frac{1}{n-1}, \frac{2}{n-1}, \dots, \frac{n-2}{n-1}, 1\} \quad (2.7)$$

Предложенную модель можно получить, как предельный переход от существующих моделей классификации физических лиц по кредитоспособности, 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 (2.8).

$$|Q^{(n)}(x) - Q(x)| < \frac{1}{n-1}, \forall x \in X. \quad (2.8)$$

При увеличении числа градаций величина градации стремится к нулю (2.9).

$$\lim_{n \rightarrow +\infty} \frac{1}{n-1} = 0. \quad (2.9)$$

Таким образом, при любом  $\varepsilon > 0$  существует число  $N_\varepsilon$ , такое что при  $n > N_\varepsilon$ ,  $|Q^{(n)}(x) - Q(x)| < \varepsilon$ . Следовательно, последовательность функций равномерно сходится к функции  $Q$  (2.10).

$$\{Q^{(n)}\} \Rightarrow Q \iff \lim_{n \rightarrow +\infty} \sup_{x \in X} |Q^{(n)}(x) - Q(x)| = 0. \quad (2.10)$$

В результате модель классификации объектов по непрерывной шкале является пределом для моделей классификации объектов по дискретной шкале при увеличении числа градаций (классов) до бесконечности. 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 обеспечивает более высокую финансовую стабильность банка при колебаниях рыночной конъюнктуры.

Стоит отметить, что в процессе реализации некоторых методов классификации физических лиц по кредитоспособности присутствует величина, варьируемая в непрерывном диапазоне, которую можно было бы использовать в основе меры для определения кредитоспособности по непрерывной шкале. Например, теория нечётких множеств основана на функциях принадлежности, которые варьируют в диапазоне от 0 до 1. При вычислении единой оценки на основе



множества характеристик так же может быть получена непрерывная мера. Однако большинство моделей и методов классификации физических лиц по кредитоспособности, в том числе, основанные на теории нечётких множеств ориентированы на получение бинарной оценки, остальные на получение дискретной оценки. Таким образом, предложенная модель классификации объектов по непрерывной шкале на примере определения кредитоспособности физических является новой.

Результат классификации полученный при использовании предложенной модели можно охарактеризовать, как индивидуальный класс кредитоспособности физического лица, в котором учтены все значимые характеристики клиента. Процентная ставка по кредиту, варьируемая в некотором диапазоне, конкретное значение которой зависит от класса кредитоспособности по непрерывной шкале так же может быть названа, как индивидуальная процентная ставка, в которой учтены все значимые характеристики клиента.

## 2.2 Модифицированный метод рандомизированных сводных показателей

Построить оценку кредитоспособности по непрерывной шкале можно на основе метода рандомизированных сводных показателей, описание которого осуществлено в соответствии с трудами [12, 39, 83, 125]. В текущем параграфе указаны места, где автором предложено модифицировать метод рандомизированных сводных показателей.

В общем, при сравнении объектов существенным аспектом является сложность оцениваемых характеристик. Весьма легко сравнить объекты по одному параметру. При наличии двух параметров явное доминирование первого объекта над вторым возможно в случае превосходства по обеим характеристикам или превосходства по одной их характеристик при условии равенства по другой. В общем случае при оценке по  $n$  характеристикам явное доминирование первого объекта  $x$  над вторым  $y$  достигается, когда по всем характеристикам первый объект не хуже второго и существует характеристика, по которой первый явно превосходит второй (2.11).

$$x_i \succcurlyeq y_i, i = 1, 2, \dots, n, \exists k \in \{1, 2, \dots, n\} : x_k \succ y_k \Rightarrow x \succ y \quad (2.11)$$

Трудность сравнения объектов при большом числе характеристик можно описать с помощью вероятности. Допустим, что вероятность доминирования первого объекта над вторым по каждой характеристике равно 50%. При оценке по  $n$  характеристикам вероятность полного доминирования первого объекта над вторым равна  $\frac{1}{2^{n-1}}$ . Таким образом, вероятность полного доминирования одного объекта над другим снижается при увеличении числа оцениваемых характеристик. При оценке по 9 характеристикам ( $n = 9$ ) вероятность доминирования первого объекта над вторым равна 0,39% (2.12).

$$P(x \succ y; n = 9) = \frac{1}{2^8} \quad (2.12)$$

В текущем параграфе рассматривается метод рандомизированных сводных показателей, с помощью которого можно на основе  $n$  исходных характеристик вычислить единую оценку. На основании полученной оценки можно упорядо-

чить объекты. Для начала рассмотрим упрощенную версию рассматриваемого метода, а именно метод сводных показателей. Последовательность действий для вычисления сводной оценки выглядит следующим образом:

1. Составить вектор характеристик объекта, физического лица, по которым будет осуществляться оценка (2.13).

$$x = (x_1, x_2, \dots, x_m) \quad (2.13)$$

2. Оценить проявление качества по каждой характеристике в диапазоне от 0 до 1 (2.14).

$$q_i : x_i \rightarrow [0,1], i = 1, 2, \dots, m \quad (2.14)$$

3. Определить весовые коэффициенты, которые задают значимость отдельных характеристик (2.15).

$$w = (w_1, w_2, \dots, w_m), w_i \geq 0, \sum_{i=1}^m w_i = 1 \quad (2.15)$$

4. Установить агрегирующую функцию, по которой будут объединены исходные характеристики с учётом значимости.

Для однозначного упорядочивания объектов агрегирующая функция должна удовлетворять условию монотонности и простейшим краевым условиям. Другими словами, чем больше оценка, тем предпочтительнее объект и при  $Q = 0$  оценка объекта является наихудшей, а при  $Q = 1$  – наилучшей. Оценка по описанному выше алгоритму может быть записана в общем виде (2.16).

$$Q^{(j)} = Q(q^{(j)}; w) = Q(q_1(x_1^{(j)}), q_2(x_2^{(j)}), \dots, q_m(x_m^{(j)}); w_1, w_2, \dots, w_m) \quad (2.16)$$

При оценке благосостояния общества используют функцию на подобии (2.17). В дальнейшем будет использовано  $\beta = 1$ , как и в ряде других работ [7, 130], преобразуя исходное выражение в функцию средневзвешенного арифметического.

$$Q^{(j)} = \begin{cases} (\sum_{i=1}^m q_i(x_i^{(j)})^\beta * w_i)^{\frac{1}{\beta}}, & \text{при } \beta \neq 0 \\ \prod_{i=1}^m q_i(x_i^{(j)})^{w_i}, & \text{при } \beta = 0 \end{cases} \quad (2.17)$$

где сумма весовых коэффициентов равна единице  $\sum_{i=1}^m w_i = 1$ , при неотрицательности каждого из них  $w_i \geq 0$ .

Оценка объекта по описанному выше алгоритму вычисляется, когда известны все составные элементы. Однако значимость весовых коэффициентов заранее неизвестна. Здесь возникает потребность использовать метод рандомизированных сводных показателей. По данному методу весовые коэффициенты рандомизируются с равномерным распределением по множеству, образованному по дополнительной информации. В результате набор весовых коэффициентов и оценка объекта становятся случайными величинами (2.18, 2.19). Для оценки случайной величины  $\tilde{Q}$  используется математическое ожидание (2.20). Для определения точности полученной оценки  $\bar{Q}$  вычисляется стандартное отклонение (2.21).

$$\tilde{w} = (\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_m) \quad (2.18)$$

$$\tilde{Q} = \tilde{Q}(q; \tilde{w}) \quad (2.19)$$

$$\bar{Q} = E(\tilde{Q}) = E(\tilde{Q}(q; \tilde{w})) \quad (2.20)$$

$$S\tilde{Q} = \sqrt{D(\tilde{Q})} = \sqrt{D(\tilde{Q}(q; \tilde{w}))} \quad (2.21)$$

Для сравнения объектов по указанному методу можно воспользоваться математическим ожиданием случайной величины. Однако случайная величина, как правило, либо с малой долей вероятности принимает значение, равное среднему, либо никогда не принимает среднего значения. При различной реализации случайной величины для весовых коэффициентов (2.18) соотношение оценок объектов может отличаться. Поэтому в дополнение к математическому ожиданию может быть использовано доминирование по вероятности. Тот факт, что объект  $j$  доминирует над объектом  $l$  на уровне значимости  $\alpha$  означает, что веро-

ятность превосходства оценки  $Q^{(j)}$  над оценкой  $Q^{(l)}$  при реализации случайной величины больше  $\alpha$  (2.22).

$$P(\tilde{Q}^{(j)} > \tilde{Q}^{(l)}) > \alpha \quad (2.22)$$

Рандомизацию сводных показателей можно алгоритмизировать следующим образом. Дополнительная информация задаёт множество  $W$  (2.23). На данном множестве распределена случайная величина  $\tilde{w} = (\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_m)$ . Математическое ожидание и стандартное отклонение данной случайной величины являются оценкой весовых коэффициентов и параметра точности, соответственно.

$$W = \{w^{(t)} = (w_1^{(t)}, w_2^{(t)}, \dots, w_m^{(t)})\} \quad (2.23)$$

Весовые коэффициенты принимают дискретные значения с заданными промежутками (2.24). Сумма вектора весовых коэффициентов равна единице. Множество всех комбинаций данных векторов можно задать как множество  $W(m, n)$  (2.25), где  $N(m, n)$  - число всех комбинаций весовых коэффициентов (2.26).

$$w_i \in w(n) = \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right\} \quad (2.24)$$

$$W(m, n) = \{w^{(t)} = (w_1^{(t)}, w_2^{(t)}, \dots, w_m^{(t)}), w^{(t)} \in w(n), \\ w_1^{(t)} + w_2^{(t)} + \dots + w_m^{(t)} = 1, t \in \{1, 2, \dots, N(m, n)\}\} \quad (2.25)$$

$$N(m, n) = \frac{(n + m - 1)!}{n!(m - 1)!} \quad (2.26)$$

Для рандомизации вектора  $w = \{w_1, w_2, \dots, w_m\}$  на множестве  $W(m, n)$  используется случайная величина  $\tilde{t}$ , равномерно распределённая на множестве допустимых комбинаций весовых коэффициентов (2.27). В результате случайная величина  $\tilde{w}$  равномерно распределена на множестве  $W(m, n)$ .

$$P(\tilde{t} = t) = \frac{1}{N(m, n)}, t \in \{1, 2, \dots, N(m, n)\} \quad (2.27)$$

Далее для полученных случайных величин вычисляется математическое ожидание (2.28) и стандартное отклонение (2.29).

$$\bar{w}_i = E(\tilde{w}_i) = \frac{1}{N(m,n)} \sum_{t=1}^{N(m,n)} w_i^{(t)} \quad (2.28)$$

$$S\tilde{w}_i = \sqrt{D\tilde{w}_i} = \sqrt{\frac{1}{N(m,n)} \sum_{t=1}^{N(m,n)} [w_i^{(t)} - \bar{w}_i]^2} = \sqrt{\frac{(m-1)(m+n)}{nm^2(m+1)}} \quad (2.29)$$

Теперь будет приведён пример вычисления оценки по указанному алгоритму, сначала без информации о значимости характеристик. Допустим, число объектов равно 3 ( $m = 3$ ), величина градации равна  $1/5$  ( $n = 5$ ). В таком случае множество  $W(3,5)$  состоит из  $N(3,5) = 21$  элемента (Табл. 2.1).

Весовые коэффициенты примерно равны  $0,333$  ( $w_1 = w_2 = w_3 \approx 0,333$ ). Стандартное отклонение для каждого весового коэффициента примерно равно  $0,298$  ( $Sw_i = 0,298, i = 1,2,3$ ). Таким образом, при отсутствии дополнительной информации о значимости атрибутов весовые коэффициенты оказываются равными (2.30), а стандартное отклонение сопоставимо с математическим ожиданием, что свидетельствует о низкой точности полученной оценки.

$$w_i = \frac{1}{m}, i = 1,2,\dots,m \quad (2.30)$$

Теперь, упорядочим 5 объектов, оцениваемых по трём атрибутам (Табл. 2.2), при отсутствии дополнительной информации о значимости характеристик. Данные характеристики необходимо нормировать, например линейной нормировкой (2.31). В результате будут получены значения, представленные в Таблице 2.3. Преобразовать исходные характеристики в оценки, варьируемые в диапазоне от 0 до 1 можно с использованием теории нечётких множеств, что будет осуществлено в следующей главе.

$$q_i^{(j)} = \frac{x_i^{(j)} - x_i^{min}}{x_i^{max} - x_i^{min}} \quad (2.31)$$

На основе значений весов и функций качества вычислены все комбинации сводной оценки  $Q^{(t)}(q^{(j)}) = Q(q^{(j)}; w^{(t)})$  (Табл. 2.4).

Таблица 2.1: Перечень весовых коэффициентов при числе атрибутов 3 и величине градации 1/5

t	$w_1^{(t)}$	$w_2^{(t)}$	$w_3^{(t)}$
1	1	0	0
2	0,8	0,2	0
3	0,6	0,4	0
4	0,4	0,6	0
5	0,2	0,8	0
6	0	1	0
7	0	0,8	0,2
8	0	0,6	0,4
9	0	0,4	0,6
10	0	0,2	0,8
11	0	0	1
12	0,2	0	0,8
13	0,4	0	0,6
14	0,6	0	0,4
15	0,8	0	0,2
16	0,2	0,2	0,6
17	0,4	0,2	0,4
18	0,6	0,2	0,2
19	0,2	0,4	0,4
20	0,4	0,4	0,2
21	0,2	0,6	0,2

Таблица 2.2: Значения исходных характеристик четырёх объектов, оцениваемых по трём атрибутам

j	$x_1^{(j)}$	$x_2^{(j)}$	$x_3^{(j)}$
1	1	2	5
2	3	4	2
3	5	1	2
4	4	3	1
5	2	3	5

На основе данных Таблицы 2.4 вычислены математическое ожидание (2.32) и стандартное отклонение (2.33) для весовых коэффициентов. В Таблице 2.5 представлены математическое ожидание ( $\bar{Q}^{(j)}$ ), стандартное отклонение ( $S\tilde{Q}^{(j)}$ ), минимальное ( $\min(\tilde{Q}^{(j)})$ ) и максимальное ( $\max(\tilde{Q}^{(j)})$ ) значения стохастической оценки.

Таблица 2.3: Нормированные значения исходных характеристик

t	$q_1^{(j)}$	$q_2^{(j)}$	$q_3^{(j)}$
1	0	0,25	1
2	0,5	0,75	0,25
3	1	0	0,25
4	0,75	1	0
5	0,25	0,5	0,75

Таблица 2.4: Значения оценок при различных весовых коэффициентах

t	$Q^{(t)}(q^{(1)})$	$Q^{(t)}(q^{(2)})$	$Q^{(t)}(q^{(3)})$	$Q^{(t)}(q^{(4)})$	$Q^{(t)}(q^{(5)})$
1	0	0,5	1	0,75	0,25
2	0,05	0,55	0,8	0,8	0,3
3	0,1	0,6	0,6	0,85	0,35
4	0,15	0,65	0,4	0,9	0,4
5	0,2	0,7	0,2	0,95	0,45
6	0,25	0,75	0	1	0,5
7	0,4	0,65	0,05	0,8	0,55
8	0,55	0,55	0,1	0,6	0,6
9	0,7	0,45	0,15	0,4	0,65
10	0,85	0,35	0,2	0,2	0,7
11	1	0,25	0,25	0	0,75
12	0,8	0,3	0,4	0,15	0,65
13	0,6	0,35	0,55	0,3	0,55
14	0,4	0,4	0,7	0,45	0,45
15	0,2	0,45	0,85	0,6	0,35
16	0,65	0,4	0,35	0,35	0,6
17	0,45	0,45	0,5	0,5	0,5
18	0,25	0,5	0,65	0,65	0,4
19	0,5	0,5	0,3	0,55	0,55
20	0,3	0,55	0,45	0,7	0,45
21	0,35	0,6	0,25	0,75	0,5

$$\bar{Q}^{(j)} = E(\tilde{Q}^{(j)}) = E(\tilde{Q}(q^{(j)}; \tilde{w})) = \frac{1}{N(m,n)} \sum_{t=1}^{N(m,n)} Q^{(t)}(q^{(j)}) \quad (2.32)$$

$$S\tilde{Q}^{(j)} = \sqrt{D\bar{Q}^{(j)}} = \sqrt{\frac{1}{N(m,n)} \sum_{t=1}^{N(m,n)} [Q^{(t)}(q^{(j)}) - \bar{Q}^{(j)}]^2} \quad (2.33)$$



Таблица 2.5: Параметры стохастической оценки

$j$	$\bar{Q}^{(j)}$	$S\tilde{Q}^{(j)}$	$\min(\tilde{Q}^{(j)})$	$\max(\tilde{Q}^{(j)})$
1	0,417	0,269	0	1
2	0,500	0,129	0,25	0,75
3	0,417	0,269	0	1
4	0,583	0,269	0	1
5	0,500	0,129	0,25	0,75

Далее вычисляются вероятности попарного доминирования объекта  $j$  над объектом  $l$  по формуле (2.34) (Табл. 2.6):

$$P(j,l) = P(\tilde{Q}^{(j)} > \tilde{Q}^{(l)}) = \frac{N\{t : Q^{(t)}(q^{(j)}) > Q^{(t)}(q^{(l)})\}}{N(m,n)}, \quad (2.34)$$

где  $N\{t : Q^{(t)}(q^{(j)}) > Q^{(t)}(q^{(l)})\}$  – число комбинаций весовых коэффициентов, при которых оценка объекта  $j$  превосходит оценку  $l$  ( $\{t : Q^{(t)}(q^{(j)}) > Q^{(t)}(q^{(l)})\}$ ). В результате деления данного числа на общее число комбинаций получается вероятность доминирования объекта  $j$  над объектом  $l$ .

Таблица 2.6: Вероятности попарного доминирования рандомизированных сводных показателей

$j \backslash l$	1	2	3	4	5
1	0,0%	28,6%	52,4%	28,6%	28,6%
2	52,4%	0,0%	52,4%	28,6%	52,4%
3	42,9%	38,1%	0,0%	28,6%	28,6%
4	71,4%	71,4%	47,6%	0,0%	52,4%
5	71,4%	47,6%	52,4%	28,6%	0,0%

Как было показано ранее отсутствие информации о значимости характеристик делает используемые характеристики равнозначными. Для придания информативности и большей практической значимости оценки объекта необходимо использовать дополнительную информацию. Допустим, что в наличии есть экспертная информация.

От экспертов весьма затруднительно получить точное значение весовых коэффициентов, потому что высока вероятность составления противоречивой системы. Противоречивая система будет получена в случае невыполнения нормирующего условия при неотрицательности весовых коэффициентов (2.35).

$$\{w_1 = 0,5, w_2 = 0,4, w_3 = 0,2, \sum_{i=1}^3 w_i \neq 1, w_i \geq 0, i = 1,2,3\} \quad (2.35)$$

Информация экспертов во многих случаях может быть представлена в менее точном виде, в виде нечисловой и неточной информации:

- нечисловая, порядковая информация задаёт попарные соотношения значимости характеристик строгими равенствами и неравенствами весовых коэффициентов (2.36);

$$OI = \{w_l > w_k, w_e = w_s, l, k, e, s \in \{1, 2, \dots, m\}\} \quad (2.36)$$

- неточная, интервальная информация задаёт диапазон варьирования случайной величины нестрогими неравенствами (2.37).

$$UI = \{0 \leq f_i \leq w_i \leq g_i \leq 1, i = 1, 2, \dots, m\} \quad (2.37)$$

Учитывая высокую степень вариативности в задании дополнительной информации по сравнению с другими экспертными методами, например методом анализа иерархий, полученные результаты обладают высоким уровнем доверия.

Неточная информация может быть так же использована для определения доминирования значимости характеристик (2.38, 2.39).

$$\begin{cases} \sum_{i=1}^3 w_i = 1 \\ 0,1 < w_1 < 0,3 \\ 0,3 < w_2 < 0,5 \end{cases} \Rightarrow w_1 < w_2 \quad (2.38)$$

$$\begin{cases} \sum_{i=1}^3 w_i = 1 \\ w_1 > 0,5 \end{cases} \Rightarrow \begin{cases} w_1 > w_2 \\ w_1 > w_3 \end{cases} \quad (2.39)$$

Далее можно определить влияние дополнительной информации на получаемый результат. Задача определения значимости характеристик сводится к вычислению характеристик случайной величины, равномерно распределённой на множестве  $W(m, n, I)$ , образованного из множества допустимых значений весовых коэффициентов с учётом ограничений, формируемых дополнительной

информацией (2.40).  $T(m, n, I)$  – номера комбинаций векторов весовых коэффициентов, удовлетворяющих условиям, формируемых дополнительной информацией.

$$W(m, n, I) = \{w^{(t)} = (w_1^{(t)}, w_2^{(t)}, \dots, w_m^{(t)}), w^{(t)} \in W(m, n), \\ w_1^{(t)} + w_2^{(t)} + \dots + w_m^{(t)} = 1, t \in T(m, n, I)\} \quad (2.40)$$

Стоит отметить, что при наличии дополнительной информации множество допустимых векторов весовых коэффициентов не больше исходного (2.41) и число допустимых векторов не больше исходного (2.42).

$$W(m, n, I) \subseteq W(m, n) \quad (2.41)$$

$$N(m, n, I) \leq N(m, n) \quad (2.42)$$

При наличии нетривиальных ограничений в дополнительной информации множество допустимых значений векторов весовых коэффициентов и число допустим векторов строго меньше исходного (2.43, 2.44).

$$W(m, n, I) \subset W(m, n) \quad (2.43)$$

$$N(m, n, I) < N(m, n) \quad (2.44)$$

Рассмотрим применение дополнительной информации на следующем примере. К предыдущей задаче добавим дополнительную информацию  $I$  (2.45). В таком случае число допустимых комбинаций весовых коэффициентов снизится более чем в 4 раза, а именно с 21 до 5 (Табл. 2.7).

$$I = \{w_1 \leq 0,2; w_2 > w_3\} \quad (2.45)$$

В случае использования дополнительной информации  $w_2 > w_3$  при наличии только характеристик  $w_2$  и  $w_3$  множество допустимых значений для вектора весовых коэффициентов находится на пересечении области выше и левее гипер-

Таблица 2.7: Допустимые комбинации весовых коэффициентов при учёте дополнительной информации

t	$w_1^{(t)}$	$w_2^{(t)}$	$w_3^{(t)}$
1	0	1	0
2	0,2	0,8	0
3	0	0,8	0,2
4	0,2	0,6	0,2
5	0	0,6	0,4

плоскости  $w_2 = w_3$  и нормирующего условия  $w_2 + w_3 = 1$ , формируя отрезок (Рис. 2.1). В рассматриваемом случае трёх весовых коэффициентов и дополнительной информации  $I$  (2.45) область допустимых значений является фигурой на плоскости. Учитывая, что весовой коэффициент  $w_1$  связан с  $w_2$  и  $w_3$  только нормирующим условием, то результирующую область можно представить на двумерном графике (Рис. 2.2). Следует помнить, что заштрихованная область на Рисунке 2.2 является плоскостью в осях  $(w_1, w_2, w_3)$ .

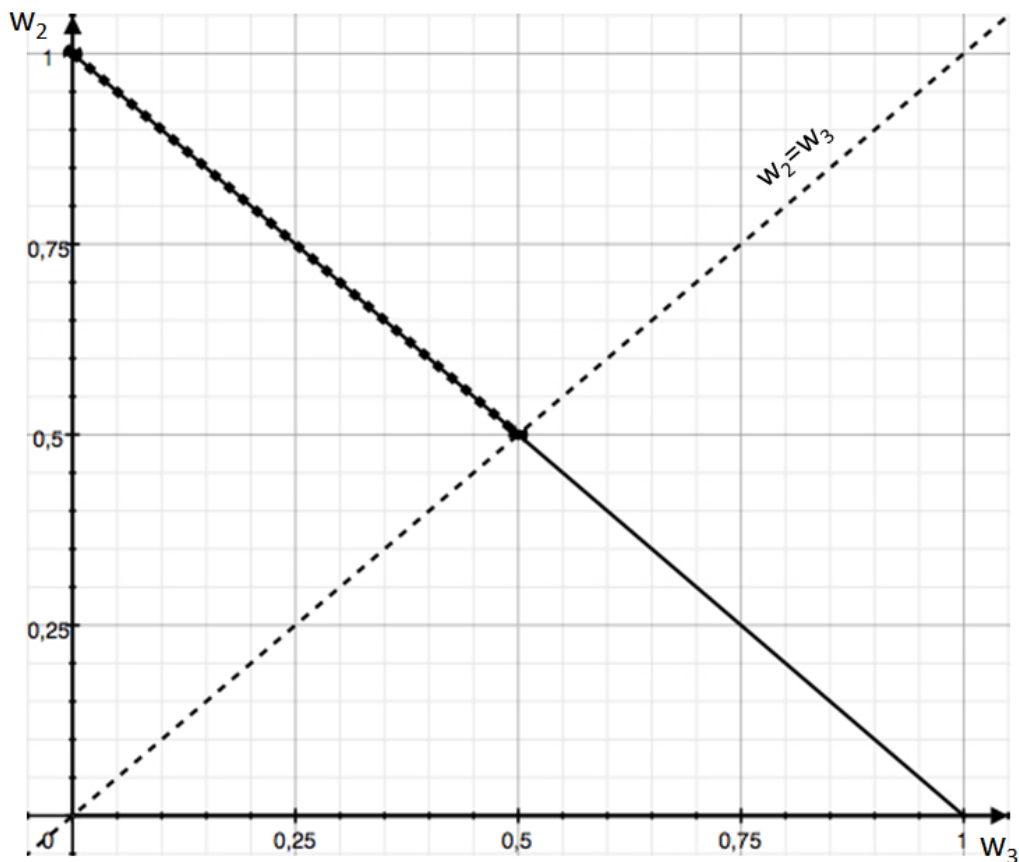


Рисунок 2.1: Область допустимых значений в осях  $(w_2, w_3)$  при ограничениях  $w_2 > w_3, w_2 + w_3 = 1$

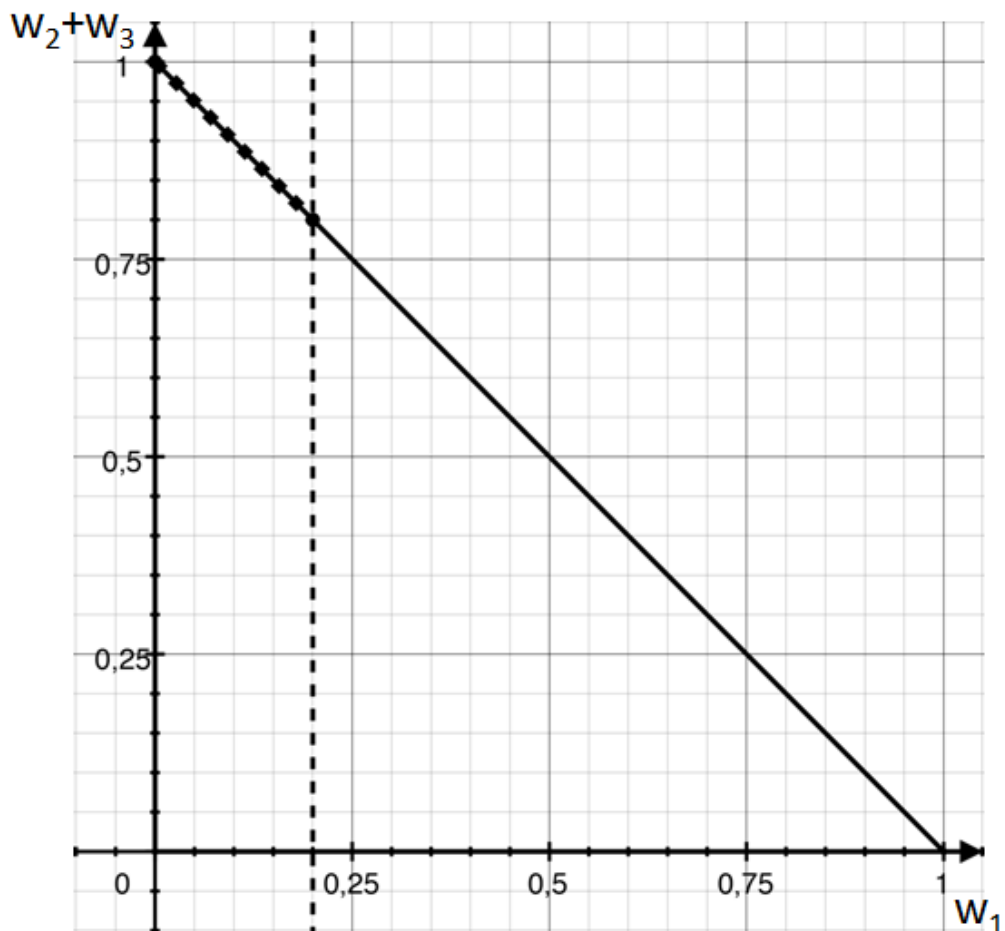


Рисунок 2.2: Область допустимых значений в осях  $(w_1, w_2 + w_3)$  при ограничениях  $w_1 \leq 0,2, w_2 > w_3, w_1 + w_2 + w_3 = 1$

В общем случае, при  $m$  невырожденных весовых коэффициентах ( $m \geq 3$ ) область допустимых значений является  $(m-1)$ -мерным многогранником. Дополнительная информация отсекает гиперплоскости от данного многогранника, формируя новое множество допустимых значений. Если дополнительная информация противоречива, то будет получено пустое множество.

Использование дополнительной информации существенно изменило характеристики случайной величины  $\tilde{w}$  (Табл. 2.8). Математическое ожидание  $\tilde{w}_1$  почти в 10 раз больше (9,5), чем для  $\tilde{w}_3$ , что означает, что значимость первой характеристики на порядок выше в искомой оценке, нежели третьей. Математическое ожидание коэффициента для первой характеристики более, чем в пять раз превосходит стандартное отклонение, что можно трактовать, как высокую точность оценки в отличие от ситуации отсутствия дополнительной информации, где данные величины сопоставимы.

Таблица 2.8: Параметры рандомизированных весовых коэффициентов при использовании дополнительной информации

i	$\bar{w}_i(I)$	$Sw_i(I)$	$\min(w_i(I))$	$\max(w_i(I))$
1	0.76	0.15	0.6	1
2	0.16	0.15	0	0.4
3	0.08	0.098	0	0.2

Также может быть вычислена вероятность доминирования одного весового коэффициента над другим (2.46).

$$p(r,s; I) = P(\tilde{w}_r(I) > \tilde{w}_s(I)) = \frac{N\{t : w_r^{(t)} > w_s^{(t)}\}}{N(m,n; I)}, \quad (2.46)$$

где  $N\{t : w_r^{(t)} > w_s^{(t)}\}$  – число элементов множества  $\{t : w_r^{(t)} > w_s^{(t)}\}$ .

Значения оценок объектов с учетом дополнительной информации посчитаны по формуле (2.47) (Табл. 2.9).

$$Q_j^{(t)}(I) = Q(q^{(j)}, w^{(t)}; I) = \sum_{i=1}^m w_i^{(t)} q_i^{(j)}, j = 1, 2, 3, 4, 5 \quad (2.47)$$

Таблица 2.9: Значения оценок при учёте дополнительной информации

t	$Q^{(t)}(q^{(1)})$	$Q^{(t)}(q^{(2)})$	$Q^{(t)}(q^{(3)})$	$Q^{(t)}(q^{(4)})$	$Q^{(t)}(q^{(5)})$
1	0,25	0,75	0	1	0,5
2	0,1	0,6	0,6	0,85	0,45
3	0,2	0,45	0,85	0,6	0,55
4	0,05	0,55	0,8	0,8	0,5
5	0	0,5	1	0,75	0,6

Математическое ожидание (2.48) и стандартное отклонение (2.49) для рандомизированных оценок представлены в Таблице 2.10.

$$\bar{Q}^{(j)}(I) = E(\tilde{Q}^{(j)}(I)) = E(\tilde{Q}(q^{(j)}; \tilde{w}(I))) = \frac{1}{N(m,n; I)} \sum_{t=1}^{N(m,n; I)} Q^{(t)}(q^{(j)}) \quad (2.48)$$

Таблица 2.10: Параметры рандомизированных оценок при учёте дополнительной информации

j	$\bar{Q}^{(j)}(I)$	$SQ^{(j)}(I)$	$\min(Q^{(j)}(I))$	$\max(Q^{(j)}(I))$
1	0.12	0.093	0	0.25
2	0.57	0.103	0.45	0.75
3	0.65	0.349	0	1
4	0.80	0.130	0.6	1
5	0.52	0.051	0.45	0.6

$$S\tilde{Q}^{(j)}(I) = \sqrt{D\bar{Q}^{(j)}(I)} = \sqrt{\frac{1}{N(m,n;I)} \sum_{t=1}^{N(m,n;I)} [Q^{(t)}(q^{(j)}) - \bar{Q}^{(j)}(I)]^2} \quad (2.49)$$

Рассматриваемые пять объектов можно однозначно ранжировать по построенным оценкам  $\bar{Q}^{(j)}(I)$ , например в порядке убывания: 4, 3, 2, 5, 1. Однако объекты 2, 3 и 5 весьма схожи по среднему значению оценки. В качестве дополнительного параметра при ранжировании вычисляется вероятность попарного доминирования при учёте дополнительной информации (2.50) (Табл. 2.11).

$$P(j,l) = P(\tilde{Q}^{(j)} > \tilde{Q}^{(l)}) = \frac{N\{t : Q^{(t)}(q^{(j)}) > Q^{(t)}(q^{(l)})\}}{N(m,n)} \quad (2.50)$$

Таблица 2.11: Вероятности попарного доминирования рандомизированных оценок при учёте дополнительной информации

j \ l	1	2	3	4	5
1	0%	0%	20%	0%	0%
2	100%	0%	20%	0%	60%
3	80%	60%	0%	40%	80%
4	100%	100%	40%	0%	100%
5	100%	40%	20%	0%	0%

Объект 3 превосходит объект 2 в 60% случаев. В 20% случаев объект 2 превосходит объект 3. В 20% случаев данные объекты равны. Следовательно объект 3 доминирует над объектом 2 по построенному показателю. Объект 2 так же доминирует над объектом 5, но с меньшей вероятностью близкой к ситуации стохастической неразличимости. Доминирование объекта 3 над объектом

5 более существенно и происходит в 80% случаев. В результате по вероятности попарного доминирования рассматриваемые объекты ранжируются таким же образом, как и по построенным оценкам  $\overline{Q}^{(j)}(I)$ . Стоит отметить, что некоторая информация о доминировании искомым оценок могла быть получена из Таблицы 2.10. Например, оценка первого объекта варьирует от 0 до 0,25, что меньше, чем нижняя грань варьирования оценки для второго, четвёртого и пятого объектов, поэтому значение вероятности доминирования в первом столбце второй, четвертой и пятой строк Таблицы 2.11 равно 100%.

При отсеве некредитоспособных клиентов высокое 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 долю потерь по ссудам на единицу выданных средств (2.51).

$$c(TA) = 1 - TA \quad (2.51)$$

В связи с этим целесообразно настроить параметры предложенной модели классификации объектов таким образом, чтобы доля потерь по ссудам принимала наименьшее значение. На основе описанного выше метода можно сформулировать оптимизационную задачу (2.52, 2.53).

$$c(TA(Q(w))) \rightarrow \min, \quad (2.52)$$

$$\left\{ \begin{array}{l} w_i - g_i \leq 0, i = 1, 2, \dots, m, \\ w_i - f_i \geq 0, i = 1, 2, \dots, m, \\ w_i - w_j > 0, \text{ при } L(i, j) = 1, \\ w_i - w_j = 0, \text{ при } E(i, j) = 1, \\ 0 \leq f_i \leq g_i \leq 1, i = 1, 2, \dots, m, \\ \sum_{i=1}^m w_i = 1, w_j \geq 0, j = 1, 2, \dots, m \end{array} \right. \quad (2.53)$$

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 коэффициентов  $w_i$  и  $w_j$  отсутствует.

В терминах метода рандомизированных сводных показателей задачу можно сформулировать следующим образом (2.54).

$$c(TA(Q(w))) \rightarrow \min, w \in OI \cap UI \cap NC, \quad (2.54)$$

где  $OI$  – нечисловая (ординальная) информация задаёт порядок доминирования весовых коэффициентов, которые определяют значимость характеристик клиентов,  $OI = \{w_l > w_k, w_e = w_s, l, k, e, s \in \{1, 2, \dots, m\}\}$ ;  $UI$  – неточная информация задаёт 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 весовых коэффициентов сформированное нормирующим условием,  $NC = \{w = (w_1, w_2, \dots, w_m) : \sum_{i=1}^m w_i = 1, w_i \geq 0\}$ .

## 2.3 Метод классификации объектов по непрерывной шкале

Постановка задачи оценки кредитоспособности состоит в следующем. Банк обладает некоторой информацией о потенциальном заёмщике. На основе данной информации банку необходимо оценить заёмщиков по уровню кредитоспособности. В текущем параграфе кредитоспособность оценена с помощью стандартной версии метода рандомизированных сводных показателей применённая в области кредитования физических лиц. Для того чтобы составить оценку кредитоспособности физического лица, необходимо выполнить следующие действия:

1. Составить вектор характеристик объекта, физического лица по имеющейся в банке информации (2.55).

$$x = (x_1, x_2, \dots, x_m) \quad (2.55)$$

2. Составить функции качеств (2.56), каждая из которых определяет степень проявления качества по отдельно взятой характеристике.

$$q_i(x_i) \in [0, 1], i = 1, 2, \dots, m \quad (2.56)$$

3. Вычислить весовые коэффициенты (2.57), определяющие значимость каждой характеристики в оценке кредитоспособности

$$w = (w_1, w_2, \dots, w_m), w_i \geq 0, \sum_{i=1}^m w_i = 1 \quad (2.57)$$

4. Определить агрегирующую функцию, по которой будет определена результирующая оценка кредитоспособности физического лица по всем  $m$  функциям качеств потенциального заёмщика (2.58).

$$Q = Q(q; w), Q : [0, 1]^m \longrightarrow [0, 1] \quad (2.58)$$

Для заёмщика  $j$  с характеристиками  $x^{(j)}$  (2.59) соответствует вектор оценок качества по каждой характеристике  $q^{(j)}$  (2.60). При заданных значениях весовых коэффициентов  $w^{(0)}$  (2.61),  $j$  – му заёмщику однозначно определена оценка кредитоспособности (2.62).

$$x^{(j)} = (x_1^{(j)}, x_2^{(j)}, \dots, x_m^{(j)}) \quad (2.59)$$

$$q^{(j)} = (q_1^{(j)}, q_2^{(j)}, \dots, q_m^{(j)}) \quad (2.60)$$

$$w^{(0)} = (w_1^{(0)}, w_2^{(0)}, \dots, w_m^{(0)}) \quad (2.61)$$

$$Q^{(j)} = Q(q^{(j)}; w^{(0)}) = Q(q_1(x_1^{(j)}), q_2(x_2^{(j)}), \dots, q_m(x_m^{(j)}); w_1^{(0)}, w_2^{(0)}, \dots, w_m^{(0)}) \quad (2.62)$$

При анализе частной информации физических лиц во многих случаях по некоторым характеристикам отсутствуют данные, что соответствует нулевому значению функции качества. Использование функции средневзвешенного геометрического сделает данных физических лиц, а так же тех, у кого функция качества хотя бы по одной характеристике равна нулю, неразличимыми, потому что оценка кредитоспособности будет равна нулю (2.63). В качестве агрегирующей функции использована функция средневзвешенного арифметического (2.64).

$$Q^{(j)} = \prod_{i=1}^m q_i(x_i^{(j)})^{w_i}, w_i \geq 0, \sum_{i=1}^m w_i = 1, \exists q_i(x_i^{(j)}) = 0 \Rightarrow Q^{(j)} = 0 \quad (2.63)$$

$$Q^{(j)} = Q(q^{(j)}; w^{(0)}) = \sum_{i=1}^m q_i(x_i^{(j)}) * w_i^{(0)}, w_i^{(0)} \geq 0, \sum_{i=1}^m w_i^{(0)} = 1 \quad (2.64)$$

Помимо получения частной информации клиентов, необходимо определить конкретное значение весовых коэффициентов. Весовые коэффициенты определяются посредством рандомизации на множестве допустимых значений. Множество допустимых значений вектора весовых коэффициентов определяется по нечисловой и неточной информации о частных характеристиках потенциального заёмщика. Нечисловая (ординальная) информация задаёт порядок доминирования значимости (весовых коэффициентов) любой пары частных характеристик заёмщиков (2.36). Неточная информация задаёт диапазон варьирования значимости частных характеристик потенциальных заёмщиков (2.37).

В результате весовые коэффициенты представлены в виде случайной величины (2.65). Точное значение весовых коэффициентов оценивается, как математическое ожидание от соответствующей случайной величины (2.66), что используется при построении оценки кредитоспособности (2.67).

$$\tilde{w} = (\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_m) \quad (2.65)$$

$$\bar{w}_i = E(\tilde{w}_i), i = 1, 2, \dots, n \quad (2.66)$$

$$Q = Q(q; \bar{w}) = \sum_{i=1}^m q_i(x_i) * \bar{w}_i, \bar{w}_i \geq 0, \sum_{i=1}^m \bar{w}_i = 1 \quad (2.67)$$

Частная информация физических лиц разнообразна и содержит как качественные, так и количественные характеристики. До сих пор не существует стандартного набора характеристик, по которому можно классифицировать физических лиц по уровню кредитоспособности [115]. Даже для фирм не существует единого стандарта, согласно которому можно определить качество финансового состояния фирмы [126]. Структура частной информации физических лиц может быть представлена в виде иерархии характеристик (Рис. 2.3), где оценку кредитоспособности составляют:

1. социальное положение;
2. трудовые показатели;
3. имущественное положение;
4. деловая репутация.

Каждый из указанных элементов описывается более подробно на третьем уровне иерархии. Социальное положение физического лица составляют возраст, семейное положение, количество иждивенцев. Трудовые показатели определяют способность клиента зарабатывать средства, что важно с точки зрения экономического содержания. Сравнивая доходность физического лица со ставкой по кредиту, можно сразу определить, что потенциальный заёмщик не способен вернуть кредит из собственных доходов за период действия кредита. Трудовые показатели физического лица составляют образование, место работы, доходы. Имущественное положение – это средства, которыми обладает заёмщик на момент взятия кредита. Имущественное положение физического лица составляют недвижимость в собственности, автотранспорт в собственности. Деловая репутация – это характеристика надёжности физического лица по информации с места работы (характеристика коллег, наличие выговоров, взысканий и т.п.), кредитных организаций и т.п. В качестве информации, которую может получить банк с места работы о любом потенциальном заёмщике, выступает тру-

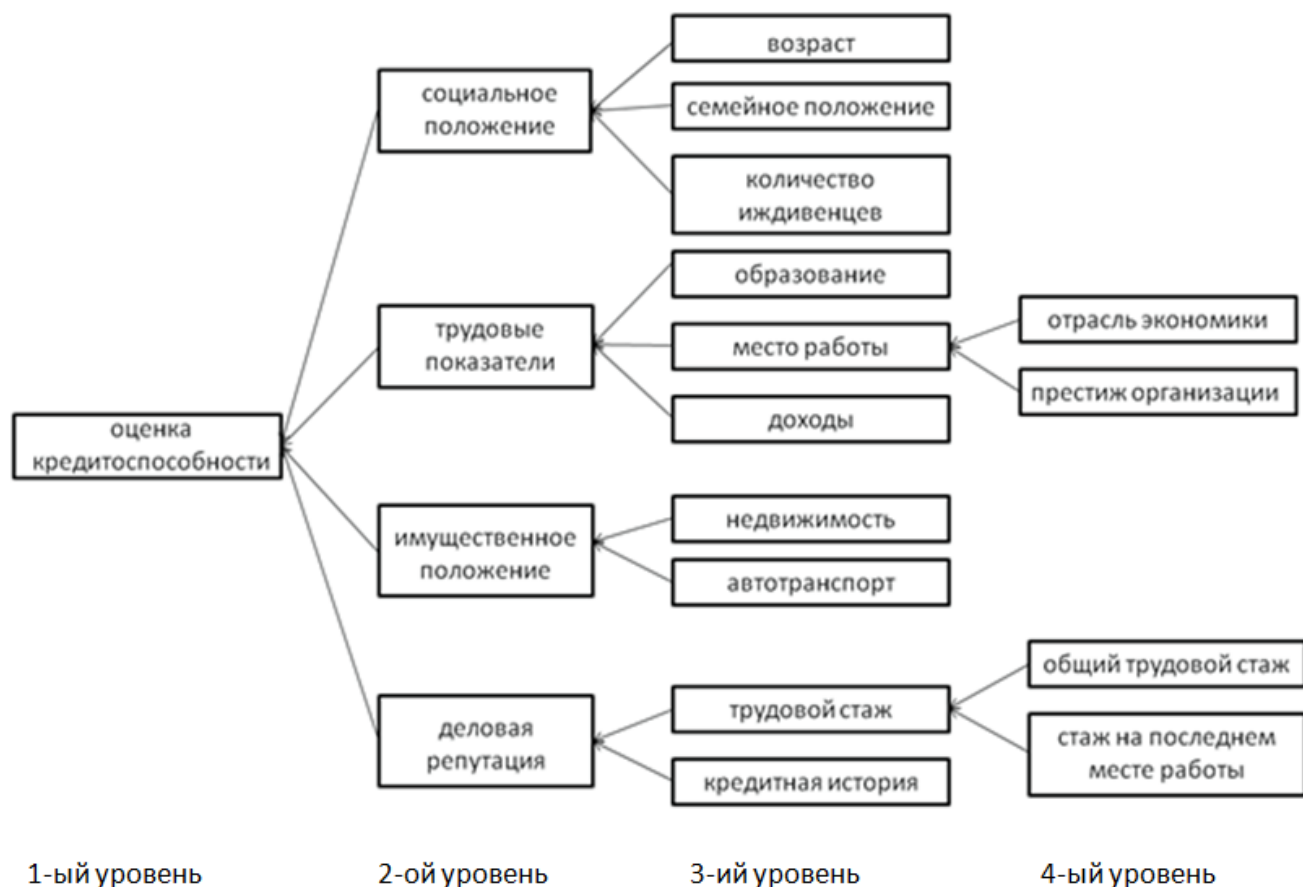


Рисунок 2.3: Частная информация физического лица, представленная в виде иерархии характеристик

довой стаж. Деловую репутацию физического лица составляют трудовой стаж, кредитная история.

Приведённая структура частной информации физического лица может быть дополнена. Например, имущественное положение может быть дополнено информацией о прочем имуществе физического лица. Характеристики третьего уровня иерархии могут быть описаны более подробно. Например, подробнее могут быть описаны характеристики места работы, приводя информацию об отрасли, в которой работает фирма и деловую репутацию фирмы. Также трудовой стаж может быть более подробно рассмотрен, как общий трудовой стаж и трудовой стаж на последнем месте работы. Взаимоотношение в домашнем хозяйстве, могут строится по различным моделям [5].

Некоторые характеристики могут быть отнесены к разным элементам иерархии. Например, образование по данной схеме отнесено к трудовым характеристикам. Однако образование может быть отнесено и к социальному положению физического лица. При анализе оценки кредитоспособности физического лица

данный факт следует учитывать, чтобы одна и та же характеристика не была многократно посчитана, что преувеличит вес (значимость) в искомой оценке. Применение предложенного метода кредитного скоринга физических лиц к рассматриваемой структуре позволяет по элементам предыдущего уровня иерархии вычислить последующие элементы.

Качественные характеристики для построения математической модели должны получить числовую оценку [2]. В приведённой структуре частной информации потенциального заемщика описание на четвёртом уровне иерархии имеют такие элементы, как место работы и трудовой стаж. Оценка места работы требует детального изучения, поэтому в рамках текущей модели потенциальные заемщики разбиты на три группы по степени надёжности места трудоустройства: высокая надёжность работодателя (ВН), средняя надёжность работодателя (СН), низкая надёжность работодателя (НН). Каждой из приведённых групп поставлено в соответствие значение функции качества: 1 - при высокой надёжности работодателя, 0,5 - при средней надёжности работодателя, 0 - при низкой надёжности работодателя. Составляющие трудового стажа измеряются следующим образом: общий трудовой стаж в годах, стаж на последнем месте работы в месяцах. Чем больше каждая из составляющих трудового стажа, тем выше функция качества для данного элемента. Оценка характеристик второго уровня иерархии построена следующим образом:

1. блок социального положения;

- Возраст измеряется в годах. Чем ближе возраст потенциального заемщика к 35 годам, тем выше значение функции качества.
- Семейное положение содержит три состояния: женат, не женат, в разводе. Каждому из указанных состояний поставлено значение функции качества 1, 0, 0,2 соответственно.
- Количество иждивенцев представлено в виде числа. Чем больше данное число, тем ниже значение функции качества.

2. блок трудовых характеристик;

- Потенциальные заемщики разделены на группы по уровню образования: ниже среднего общего (ниже СО), на уровне среднего общего (СО), на уровне начального профессионального (НП), на уровне среднего профессионального (СП), на уровне высшего профессионального (ВП) и послевузовского профессионального (ПП), для каждой из которых функция качества принимает значение 0, 0,2, 0,4, 0,6, 0,8 и 1 соответственно.
  - По степени надёжности места трудоустройства потенциальные заемщики разбиты на группы с высокой надёжностью работодателя (ВН), средней надёжностью работодателя (СН) и низкой надёжностью работодателя (НН), для каждой из которых функция качества принимает значение 1, 0,5 и 0 соответственно.
  - Среднемесячный доход потенциального заемщика измеряется в денежном выражении (в рублях). Чем больше средний доход, тем выше значение функции качества.
3. В блоке имущественного положения недвижимость и автотранспорт в собственности потенциального заемщика представлены в стоимостном выражении (в рублях). Чем выше стоимость имущества, тем выше значение функции качества.
  4. В блоке деловой репутации трудовой стаж определяется по общему трудовому стажу (измеряется в годах) и стажу на последнем рабочем месте (измеряется в месяцах). Чем выше общий трудовой стаж и стаж на последнем рабочем месте, тем выше значение функции качества для соответствующей характеристики. Помимо этого, в данном блоке потенциальные заемщики разбиты на группы с хорошей, удовлетворительной и плохой кредитной историей, для каждой из которых функция качества принимает значение 1, 0,5 и 0 соответственно. Отсутствие кредитной истории приравнивается к плохой кредитной истории.

По каждой характеристике физического лица можно провести отдельное исследование по выявлению функции качества и значимости для оценки кредитоспособности. Например, в основу оценки качества образования может быть

положена международная стандартная классификация образования [15], или иная классификация с использованием данных, приведённых в работе на основе анализа 146 стран за период с 1950 по 2010 годы [42].

Далее показано, как можно составить оценку кредитоспособности физического лица по составленному методу. Начальные условия состоят в следующем. Допустим, 10 физических лиц подали заявку на кредит размером в 100 тысяч рублей на полгода. Заёмщики предоставили в банк частную информацию. Задача банка состоит в том, чтобы классифицировать заёмщиков по уровню кредитоспособности с целью определения дальнейших условий взаимодействия с данными клиентами. Частная информация физических лиц в структурированном виде представлена в Таблицах 2.12 и 2.13. Таблица 2.12 содержит показатели, составляющие социальное положение и трудовые показатели физического лица. Таблица 2.13 содержит показатели, составляющие имущественное положение и деловую репутацию физического лица. На следующем этапе необходимо определить функцию качества для каждой характеристики. Данный этап был подробно расписан для каждой характеристике в предыдущем разделе.

Таблица 2.12: Частная информация физических лиц (часть 1)

	социальное положение			трудовые показатели		
	возраст	семейное положение	количество иждивенцев	образование	надёжность работодателя	доход
1	27	женат	0	ПП	высокая	30
2	21	не женат	0	НП	низкая	20
3	45	в разводе	1	СО	низкая	15
4	60	женат	2	НП	средняя	25
5	34	в разводе	2	СП	высокая	35
6	30	женат	0	ВП	средняя	40
7	41	женат	1	СП	средняя	30
8	20	не женат	0	ниже СО	низкая	12
9	52	в разводе	3	ВП	средняя	35
10	26	женат	1	СП	низкая	20

На третьем этапе алгоритма необходимо определить значение весовых коэффициентов. Данная информация определяет множество допустимых значений, по которому будет происходить рандомизация и вычисление итогового значения весовых коэффициентов. Множество допустимых значений может быть постро-



Таблица 2.13: Частная информация физических лиц (часть 2)

	имущество, тыс. руб.		деловая репутация		
	недвижимость	автотранспорт	кредитная история	трудовой стаж, лет общий	на посл. месте
1	600	0	хор.	3	3
2	0	0	пл.	1	0,8
3	500	600	уд.	25	2,3
4	1000	150	уд.	35	5
5	1200	0	хор.	24	1
6	2000	800	пл.	8	2,3
7	1500	600	хор.	21	0,7
8	0	0	пл.	4	0,3
9	800	300	уд.	30	3,8
10	0	200	хор.	5	0,5

ено с помощью математических моделей, социологического опроса, экспертного мнения и так далее, например:

- статистически банк установил, что образование ( $w_{21}$ ) и надёжность работодателя ( $w_{22}$ ) физического лица, подающего заявку на кредит, имеют одинаковый вес в оценке кредитоспособности ( $w_{21} = w_{22}$ );
- социологический опрос показал, что супружество ( $w_{12}$ ) не столь значимо для производительности физического лица, как возраст клиента ( $w_{11}$ ) ( $w_{11} > w_{12}$ );
- эксперт установил, что стоимость автотранспорта физического лица ( $w_{32}$ ) представляет больший интерес (имеет больше вес в итоговой оценке кредитоспособности) для банка за счёт более высокой ликвидности по сравнению с недвижимостью ( $w_{31}$ ) ( $w_{32} > w_{31}$ );
- банк вычислил, что для стабильности прогноза доходов клиента стаж на последнем рабочем ( $w_{422}$ ) месте важнее (имеет больше вес в итоговой оценке кредитоспособности) общего трудового стажа ( $w_{421}$ ) ( $w_{421} > w_{422}$ ).

Информация о значимости характеристик представлена в Таблице 2.14.

Значение весовых коэффициентов получено путём рандомизации на множестве допустимых значений образованного дополнительной информацией и вычисления математического ожидания, полученной случайной величины.

Таблица 2.14: Информация о значимости характеристик

Название	Условие
Социальное положение	$w_{11} > w_{12}$ $w_{13} \leq 0,3$
трудо- вые показатели	$w_{21} = w_{22}$ $w_{23} \geq 0,5$
Имущество	$w_{31} < w_{32}$ $w_{31} \geq 0,2$
Деловая репутация	$w_{421} < w_{422}$ $w_{41} \geq 0,4$
Оценка кредитоспособности	$w_2 > w_1$ $w_2 > w_3$ $w_4 > w_1$ $w_4 > w_3$ $w_3 \geq 0,2$

Далее посчитаны значения показателей второго уровня иерархии. На основе значения весовых коэффициентов и показателей второго уровня иерархии вычислено значение оценки кредитоспособности для каждого клиента (Табл. 2.15), которые округлены до четвёртого знака после запятой.

Таблица 2.15: Оценки кредитоспособности

Клиент	Оценка кредитоспособности
7	0,7913
5	0,7384
1	0,7147
6	0,7022
9	0,6186
4	0,5091
10	0,5032
3	0,4047
2	0,1306
8	0,0204

При необходимости более детальной оценки кредитоспособности можно увеличить количество знаков после запятой. Упорядочим клиентов по убыванию оценки кредитоспособности. Построенная оценка может быть использована бан-

ком для определения дальнейших условий взаимодействия с потенциальным заёмщиком следующим образом:

1. Отказать в выдаче кредита клиентам № 8, № 2 и № 3, как обладателям самого низкого уровня кредитоспособности. Остальным клиентам выдать кредит по единой процентной ставке.
2. Отказать в выдаче кредита только клиентам № 8 и № 2. Остальным заёмщикам дифференцировать процентную ставку по кредиту в соответствии с оценкой кредитоспособности. Клиенту № 7 предоставить наименьшую процентную ставку среди рассматриваемых физических лиц, для клиента № 5 процентную ставку чуть выше и так далее. Клиенту № 3 предоставить наибольшую процентную ставку.

Стоит отметить, что искомая оценка при использовании предложенного метода может являться многоуровневой, где результирующая оценка на одном этапе является исходной характеристикой на следующем. Таким образом, данный метод допускает представление частной информации заемщика в виде иерархии, что упрощает анализ данных.

Для построения оценки кредитоспособности физических лиц по предложенному методу банку необходимо собрать частную информацию клиентов, структурировать данную информацию, каждому элементу данной структуры определить значение функции качества и значимость в искомой оценке. Например, физическому лицу с возрастом 50 лет, при выдаче кредитной карты со стандартным кредитным лимитом будет соответствовать более высокое значение функции качества по данной характеристике, чем при ипотечном кредите сроком на 20 лет.

## Результаты главы

В данной главе разработаны модель и метод классификации объектов по непрерывной шкале, примененные к классификации физических лиц по кредитоспособности. Разработанный можно назвать модифицированным методом рандомизированных сводных показателей, в котором применяются методы теории нечётких множеств.

Использование построенного метода позволяет получить следующие преимущества:

1. результирующая оценка варьирует непрерывно в диапазоне от 0 до 1;
2. возможность использования дополнительной информации при определении значимости характеристик потенциального заемщика<sup>1</sup>.

Число, которое отражает класс кредитоспособности и принимает значения в диапазоне от 0 до 1, содержит обработанную частную информацию потенциального заёмщика в форме, которая, в отличие от бинарной оценки, позволяет банку принимать больше обоснованных управленческих решений, например:

- дифференцировать процентную ставку по непрерывной шкале в соответствии с полученной индивидуальной оценкой кредитоспособности.
- усреднить оценку по непрерывной шкале и классифицировать заёмщиков на кредитоспособных и некредитоспособных, выдавая кредит только кредитоспособным по единой процентной ставке;
- усреднить оценку для классификации по любой дискретной шкале.

Информация о значимости характеристик для осуществления классификации физических лиц по кредитоспособности может включать различные источники: анализ математических моделей, результаты социологических исследований, экспертное мнение и другие. Кроме того, возможность учёта точной, ординальной и интервальной информации при определении значимости характеристик потенциального заёмщика позволяет корректировать искомую оценку экспертами без математического образования (например, психолога). Например, эксперт может указать, что трудовые показатели значимее деловой репутации, что будет иметь соответствующий результат в оценке кредитоспособности.

---

<sup>1</sup> Данная информация может содержать, как точные значения о значимости характеристик, а так же иметь форму ординальной и интервальной информации

## Глава 3.

# Применение метода классификации объектов по непрерывной шкале при кредитовании физических лиц

В текущей главе разработанный модифицированный метод рандомизированных сводных показателей апробирован на реальных данных, по базе кредитных историй. На основе данного метода разработана система поддержки принятия решений по выдаче кредита. Составлена схема действий по внедрению разработанного метода и СППР в банк. Рассмотрены особенности ценообразования кредита при использовании разработанного метода.

## 3.1 Классификация физических лиц по базе кредитных историй

В Российской Федерации с 2005 года действует закон “О кредитных историях” [28]. По данному закону банки обязаны передавать информацию о выплатах физических лиц по кредиту в бюро кредитных историй [24]. Однако кредитные истории запрещено передавать третьим лицам без разрешения правообладателя информации (рассматриваемого физического лица), что затрудняет апробацию методов классификации. Для апробации разработанного метода использована общедоступная база кредитных историй<sup>1</sup>. В текущей работе проанализировано 1000 кредитных историй, из них промежуточные вычисления для первых

---

<sup>1</sup>источник: хранилище данных для машинного обучения [UCI Mashine Learning Repository], [электронный ресурс] режим доступа [https://archive.ics.uci.edu/ml/datasets/Statlog+\(German+Credit+Data\)](https://archive.ics.uci.edu/ml/datasets/Statlog+(German+Credit+Data)) (14.10.2015).

100 историй приведены подробно в Приложении А. Результаты классификации клиентов для тестового множества клиентов также приведены в приложении.

В используемой базе кредитных историй представлено 20 характеристик для физических лиц. Из них 7 количественных (числовых) и 13 качественных (категориальных) характеристик. Некоторые из используемых характеристик имеют следующее описание<sup>2</sup>:

A1 – Баланс счёта до востребования:

A11 – отрицательное значение счёта;

A12 – от 0 до 100 тыс. руб.;

A13 – более 100 тыс. руб.;

A14 – отсутствие счёта.

A6 – Сбережения, в том числе и облигации:

A61 – до 50 тыс. руб.;

A62 – от 50 тыс. руб. до 250 тыс. руб.;

A63 – от 250 тыс. руб. до 500 тыс. руб.;

A64 – более 500 тыс. руб.;

A65 – отсутствуют данные или отсутствуют сбережения.

A7 – Продолжительность трудоустройства на последнем рабочем месте:

A71 – не трудоустроен;

A72 – до 1 года;

A73 – от 1 года до 4 лет;

A74 – от 4 до 7 лет;

A75 – более 7 лет.

A11 – Продолжительность проживания в нынешних апартаментах.

A12 – Собственность:

---

<sup>2</sup>Полный список приведён в Приложении А.2. Допустим, что 1 у.е. = 500 рублей

A121 – жилая площадь;

A122 – отсутствует жилищная площадь, но есть счёт в строительной сберегательной кассе;

A123 – отсутствуют первые два пункта, но есть автомобиль;

A124 – отсутствует информация или нет собственности.

A13 – Возраст, измеряется в годах.

A16 – Число действующих кредитов в данном банке.

При построении оценки кредитоспособности предложено убрать характеристики, связанные с параметрами кредитной сделки, то есть характеристики № 2, № 4 и № 5. В таком случае будет построена универсальная, независимая от параметров кредитной сделки оценка кредитоспособности физического лица. В дальнейшем следует выразить зависимость кредитного риска от параметров кредита (срока кредитования, объёма кредита и др.) и от полученной оценки кредитоспособности. При подстановке конкретного значения оценки кредитоспособности будет вычислена линия уровня, отображающая все возможные варианты условий кредитования (процентная ставка, объем кредита и др.), доступные для рассматриваемого потенциального заёмщика.

При построении иерархии показателей физического лица, по которой составляется оценка кредитоспособности, необходимо помнить ограниченность восприятия экспертов и возможности статистических методов. Кратковременная память человека позволяет одновременно оценивать 6-7 характеристик, что накладывает ограничение на размер разбиения элементов иерархии. Подобным образом в статистических методах известно понятие “проблемы размерности”, которое заключается в том, что при увеличении числа оцениваемых факторов сложность вычислений увеличивается многократно, и с некоторого значения ресурсов оказывается недостаточно, чтобы своевременно решить задачу. Также существует принцип несовместимости при управлении системой, который заключается в том, что по мере возрастания сложности системы, наша способность формулировать точные и при этом осмысленные утверждения о ее поведении уменьшается вплоть до некоторого порога, за пределами которого точность и смысл становятся практически взаимоисключающими характеристиками [8].

Подобное ограничение на число градаций можно расширить, построив иерархию показателей, где каждый элемент разбивается не более чем на 7 элементов. В таком случае элемент на одном уровне имеет разбиение на 7 подэлементов, на втором 49, а на третьем 343. Следует ограничиться разбиением элемента не более, чем на 7 подэлементов для построения адекватной оценки. Характеристики физических лиц по рассматриваемой базе представлены в виде иерархии (Рис. 3.1).

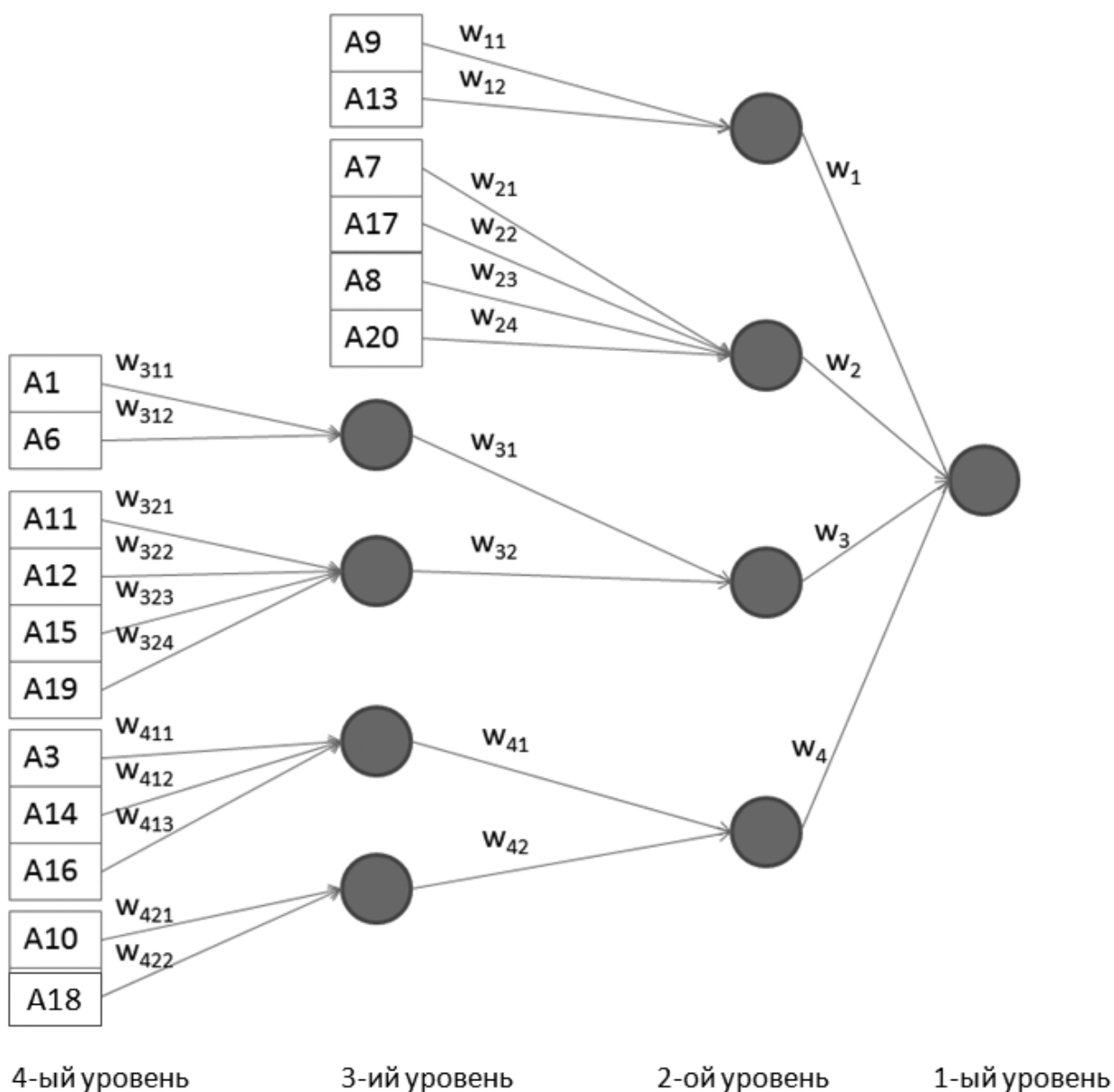


Рисунок 3.1: Иерархия характеристик физического лица для оценки кредитоспособности

Структурировав частную информацию физических лиц, оценка кредитоспособности подразделяется на следующие составные элементы:



1. социальное положение (A9, A13);
2. трудовые показатели (A7, A17, A8, A20);
3. имущество (ликвидная собственность (A1, A6), неликвидная собственность и сопутствующие показатели (A11, A12, A15, A19));
4. деловая репутация (кредитная история (A3, A14, A16), гаранты (A10, A18)).

Характеристики физических лиц необходимо нормировать. Отсутствие информации приравнивается к отрицательной информации<sup>3</sup>. Так в первой характеристике присутствует пункт A11 – отрицательный баланс счёта до востребования и пункт A14 – отсутствие счёта. Физическое лицо может иметь счёт, но скрывать это, так как баланс отрицательный, поэтому характеристике A14, как и A11 соответствует нулевое значение в модели, не увеличивая, тем самым оценку кредитоспособности. Аналогично по характеристике № 3 факт того, что клиент берёт кредит в первый раз в данном банке может означать, что предыдущий банк отказывает в дальнейшей выдаче кредита или иные мошеннические схемы, связанные с подделкой документов, поэтому при действии пункта A30 оценка по данной характеристике равна 0.

При вычислении оценки кредитоспособности все характеристики нормированы по шкале [0,1]. Причём более высокое нормированное значение по отдельной характеристике соответствует более высокому значению оценки кредитоспособности. В таком случае соотношение характеристик полностью определяется весовыми коэффициентами. По количественным характеристикам, за исключением характеристик № 13 и № 16, более высокому значению атрибута соответствует более высокое значение нормированной оценки. По характеристике № 16, чем меньше значение атрибута, тем больше значение нормированной оценки. По характеристике № 13 – возраст, значению 35 соответствует нормированная оценка 1. При отклонении от данной величины в обе стороны на величину до 15, оценка линейно уменьшается до 0, и при большем отклонении остается неизменной. По характеристикам № 19 и № 20, принимающим по два

---

<sup>3</sup>Необходимость данного положения более подробно рассматривается в параграфе 3.3.

возможных значения, оценка пунктов A191, A202 равна единице, как информация, увеличивающая оценку кредитоспособности, а оценка пунктов A192 и A201 равна 0.

В характеристике № 6 использованы граничные значения для вычисления оценки каждого пункта. Пункты A64 и A65 являются наилучшим и наихудшим значениями характеристики соответственно. Оценка для пункта A64 равна единице, а для пункта A65 равна 0. Для пунктов A61, A62 и A63 оценка равна среднему от граничных значений, нормированных на минимальную величину пятой категории. В результате  $A61=0,05$ ;  $A62=0,3$ ;  $A63=0,75$ . Аналогично в характеристике № 7 пункты A72, A73 и A74 получают оценку в соответствии со средним значением граничных условий, делённые на максимальное значение последнего пункта (7 лет). Нормированные значения атрибутов для первых 100 записей кредитных историй рассматриваемой базы представлены в Приложении А.3. Оценка качественных характеристик представлена в Таблице 3.1.

На текущем этапе уже можно построить оценку кредитоспособности. Однако в полученной оценке характеристики одного элемента будут равнозначными, то есть иметь одинаковые весовые коэффициенты при агрегировании. Поэтому целесообразно задать дополнительную информацию о соотношении характеристик (Табл. 3.2).

Основная дополнительная информация для построения оценки кредитоспособности в данном примере сформулирована следующим образом:

- Возраст для банка представляет более наглядную информацию о заёмщике, чем пол и семейное положение.
- Сбережения с учетом облигаций больше характеризует долгосрочную надёжность заёмщика, чем статус счёта до востребования. Заёмщик может попытаться обмануть банк счётом до востребования.
- Ликвидная собственность более выгодна для банка из-за высокой скорости реализации с низкими издержками по сравнению с неликвидной собственностью. Однако для заёмщика в случае наступления банкротства так же проще реализовать и увести ликвидную собственность во избежание ареста имущества при возмещении долга в банк. Таким образом при построе-

Таблица 3.1: Оценка качественных характеристик

Характеристика	Значение	Оценка	Характеристика	Значение	Оценка	
A1	A11	0	A10	A101	0	
	A12	0,5		A102	0,8	
	A13	1		A103	1	
	A14	0				
A3	A30	0	A12	A121	1	
	A31	1		A122	0,6	
	A32	0,6		A123	0,4	
	A33	0,2		A124	0	
	A34	0				
A6	A61	0,05	A14	A141	0	
	A62	0,3		A142	0,5	
	A63	0,75		A143	1	
	A64	1				
	A65	0				
A7	A71	0	A15	A151	0	
	A72	0,143		A152	1	
	A73	0,357		A153	0,7	
	A74	0,786				
	A75	1				
A9	A91	0,6	A17	A171	0	
	A92	0,8		A172	0,2	
	A93	0,2		A173	0,6	
	A94	1		A174	1	
	A95	0				
A19	A191	0	A20	A201	0	
	A192	1		A202	1	

Таблица 3.2: Дополнительная информация о соотношении характеристик характеристик третьего и четвёртого уровней иерархии

$w_{11} > w_{12}$	$w_{311} \geq 0,2$
$w_{22} > w_{23}$	$w_{31} = w_{32}$
$w_{23} > w_{21}$	$w_{412} > w_{411}$
$w_{21} \geq 0,1$	$w_{411} > w_{413}$
$w_{24} \leq 0,2$	$w_{411} \geq 0,33$
$w_{322} > w_{321}$	$w_{421} < w_{422}$
$w_{321} > w_{323}$	$w_{421} \geq 0,3$
$w_{324} \leq 0,1$	$w_{41} > w_{42}$
$w_{312} > w_{311}$	$w_{42} \geq 0,3$

нии оценки кредитоспособности ликвидная собственность и неликвидная собственность имеют равную значимость.

По располагаемой информации о значимости весовых коэффициентов необходимо вычислить их непосредственное значение. Для вычисления оценки кредитоспособности составлен код в программной среде Wolfram Mathematica 8. 1000 рассматриваемых кредитных историй представлены в виде набора списков  $H[i], i = 1, 2, \dots, 1000$ . Сначала качественные характеристики списков заменяются в соответствии с Таблицей 3.1 с помощью условного цикла, который задаётся функциями “Do” и “Replace all” (/.). По формуле (3.1) произведена замена символов качественной переменной A1 на нормированные значения. Для осуществления полной замены необходимо, через знак “/.” записать все необходимые замены по Таблице 3.1.

$$Do[NH[i] = H[i]/.A11 - > 0/.A12 - > 0,5/.A13 - > 1/.A14 - > 0, \{i, 1, 1000\}] \quad (3.1)$$

Количественные характеристики нормируются с помощью линейного преобразования. Однако, характеристику возраст не удастся нормировать стандартной линейной нормировкой, потому что зависимость оценки кредитоспособности не монотонная. Для вычисления оценки характеристики возраст использована теория нечётких множеств. Задано 3 нечётких множества для переменной возраст: “молодой”, “среднего возраста”, “пожилой”, со следующими функциями принадлежности (3.2-3.4):

$$\phi_Y(x) = \begin{cases} 1, & \text{при } x \in [18, 20] \\ \frac{40-x}{20}, & \text{при } x \in (20, 40] \\ 0, & \text{при } x \in (40, 75] \end{cases} \quad (3.2)$$

$$\phi_M(x) = \begin{cases} 0, & \text{при } x \in [18, 20] \cup (60, 75] \\ \frac{x-20}{20}, & \text{при } x \in (20, 40] \\ \frac{60-x}{20}, & \text{при } x \in (40, 60] \end{cases} \quad (3.3)$$

$$\phi_O(x) = \begin{cases} 0, & \text{при } x \in [18, 40] \\ \frac{x-40}{20}, & \text{при } x \in (40, 60] \\ 1, & \text{при } x \in (60, 75] \end{cases} \quad (3.4)$$

Результирующей переменной соответствует два нечетких множества S и L со следующими функциями принадлежности (3.5, 3.6).

$$\phi_S(y) = 1 - y, \quad y \in [0,1] \quad (3.5)$$

$$\phi_L(y) = y, \quad y \in [0,1] \quad (3.6)$$

База правил следующая:

- если возраст юный ( $x = Y$ ) или преклонный ( $x = O$ ), то оценка малая ( $y = S$ );
- если возраст средний ( $x = M$ ), то оценка большая ( $y = L$ ).

Для вычисления результирующего значения использована импликация Мамдами с максиминным методом получения результирующей функции принадлежности, где дефазификация осуществлена с помощью метода центра тяжести [22]. По методу центра тяжести момент результирующей функции принадлежности делится на площадь под графиком данной функции (3.7).

$$y_c = \frac{\int y \phi_{res}(y) dy}{\int \phi_{res}(y) dy} \quad (3.7)$$

Код для оценки характеристики возраст выглядит следующим образом (3.8).

```

listy = {}; Do[c = listphi(x)[[i]];
y = (Integrate[(1 - c)z, {z, 0, 1 - c}] + Integrate[z2, {z, 1 - c, c}] +
+ Integrate[c * z, {z, c, 1}]) / (Integrate[1 - c, {z, 0, 1 - c}] +
+ Integrate[z, {z, 1 - c, c}] + Integrate[c, {z, c, 1}]);
listy = Append[listy, y], {i, 1, 1000}]

```

(3.8)

Стоит отметить, что по методу центра тяжести итоговая оценка варьирует в диапазоне  $[\frac{1}{3}, \frac{2}{3}]$ , поэтому к полученным значениям ( $z$ ) применено линейное преобразование (3.9), которое осуществлено с применением соответствующего кода 3.10.

$$f : y \rightarrow [0,1], f(y) = 3y - 1 \quad (3.9)$$

$$Do[listy[[i]] = listy[[i]] * 3 - 1, \{i, 1, 1000\}] \quad (3.10)$$

Код для вычисления весовых коэффициентов третьего и четвёртого уровней иерархии представлен в Приложении (А.4).

Для разработанной, модифицированной версии метода потребуются только средние значения для коэффициентов третьего и четвёртого уровней. Далее множество кредитных историй разделено на две группы по 500 наблюдений в каждой. По первой группе кредитных историй будет подбираться набор весовых коэффициентов оптимизирующий значение общей точности классификатора (Total Accuracy) при ограничениях представленных дополнительной информацией  $I$ . По второй группе кредитных историй оптимальные значения весовых коэффициентов тестируются.

Далее решается сформулированная ранее оптимизационная задача минимизации доли потерь по ссудам при ограничениях сформированных дополнительной информацией и нормирующим условием (2.52, 2.53).

Целевая функция весьма нетривиального вида (дискретная функция) для оптимизационной задачи, что существенно усложняет решение. Более того, да-

же если целевая функция будет линейной, область допустимых значений имеет такой вид, что симплекс-метод не применим, потому что в базовой версии метода рандомизированных сводных показателей нечисловая информация не может быть представлена в виде нестрогого неравенства весовых коэффициентов [39,83,125]. В связи с этим решение данной задачи осуществлено с помощью инструментария метода рандомизированных сводных показателей.

Множество допустимых значений весовых коэффициентов в решаемой задаче формируются с использованием дополнительной информации  $I$  (3.11):

$$I = \{w_1 > w_3, w_2 > w_4, w_4 > w_1, w_1 \geq 0,1, w_2 \leq 0,5, w_4 \geq 0,3\} \quad (3.11)$$

С учетом дополнительной информации ограничения в решаемой оптимизационной задаче имеют следующий вид (3.12).

$$\begin{cases} w_1 - 0,1 \leq 0, w_4 - 0,3 \leq 0, w_2 - 0,5 \geq 0, \\ w_1 - w_3 > 0, w_2 - w_4 > 0, w_4 - w_1 > 0, \\ \sum_{i=1}^4 w_i = 1, w_j \geq 0, j = 1, 2, 3, 4 \end{cases} \quad (3.12)$$

Для вычисления весовых коэффициентов, которые оптимизируют значение целевой функции сначала составляются списки допустимых комбинаций весовых коэффициентов (3.13). Число комбинаций весовых коэффициентов равно  $m$  (3.14).

$$\begin{aligned} &ListW[1] = \{\}; ListW[2] = \{\}; ListW[3] = \{\}; ListW[4] = \{\}; \\ &Do[If[W1 + W2 + W3 + W4 == 1, If[W2 > W3, If[W2 > W1, \\ &If[W2 > W4, If[W4 > W1, ListW[1] = Prepend[ListW[1], W1]; \\ &ListW[2] = Prepend[ListW[2], W2]; ListW[3] = Prepend[ListW[3], W3]; \\ &ListW[4] = Prepend[ListW[4], W4], 0], 0], 0], 0], 0], \\ &\{W1, 0.1, 1, 0.01\}, \{W2, 0, 0.5, 0.01\}, \{W3, 0, 0.2, 0.01\}, \{W4, 0.3, 1, 0.01\}] \end{aligned} \quad (3.13)$$

$$m = \text{Length}[\text{ListW}[1]] \quad (3.14)$$

Затем создаётся матрица для вычисления кредитоспособности, где каждая строка соответствует своему набору весовых коэффициентов, которая заполняется значениями кредитоспособности (3.15).

$$\begin{aligned} \text{ListQtr} &= \text{Table}[0, \{j, 1, m\}, \{i, 1, 500\}]; \\ \text{Do}[\text{ListQtr}[[j, i]] &= \text{ListLevel2}[[i, 1]] * \text{ListW}[1][[j]] + \\ &\text{ListLevel2}[[i, 2]] * \text{ListW}[2][[j]] + \\ &\text{ListLevel2}[[i, 3]] * \text{ListW}[3][[j]] + \\ &\text{ListLevel2}[[i, 4]] * \text{ListW}[4][[j]], \{i, 1, 500\}, \{j, 1, m\}]; \end{aligned} \quad (3.15)$$

Далее производятся вычисления для осуществления бинарной классификации (3.16).

$$\begin{aligned} \text{BClassQ} &= \text{Table}[1, \{j, 1, m\}, \{i, 1, 500\}]; \\ \text{OrderListQ} &= \text{Table}[0, \{j, 1, m\}, \{i, 1, 136\}]; \\ \text{Do}[\text{OrderListQ}[[j, i]] &= \text{Ordering}[\text{ListQtr}[[j]]][[i]], \\ &\{i, 1, 136\}, \{j, 1, m\}]; \\ \text{Do}[\text{BClassQ}[[j, i]] &= 0, \{j, 1, m\}, \{i, \text{OrderListQ}[[j]]\}]; \end{aligned} \quad (3.16)$$

Далее вычисляется показатель общей точности (3.17).



$$\begin{aligned}
ListTP &= Table[0, \{i, 1, m\}]; ListTN = Table[0, \{i, 1, m\}]; \\
ListTA &= Table[0, \{i, 1, m\}]; ListTPO = Table[0, \{j, 1, m\}, \\
&\{i, 1, 500\}]; ListTNO = Table[0, \{j, 1, m\}, \{i, 1, 500\}]; \\
Do[If[BClassQ[[j, i]] + Cr[[i]] == 2, \\
ListTPO[[j, i]] = 1, 0], \{i, 1, 500\}, \{j, 1, m\}]; \\
Do[If[BClassQ[[j, i]] + Cr[[i]] == 0, \\
ListTNO[[j, i]] = 1, 0], \{i, 1, 500\}, \{j, 1, m\}]; \\
Do[ListTA[[j]] = Mean[ListTPO[[j]]] + Mean[ListTNO[[j]]], \{j, 1, m\}];
\end{aligned} \tag{3.17}$$

Теперь можно получить максимальное значение целевой функции и номер набора весовых коэффициентов, где достигается максимальное значение, а также значение весовых коэффициентов можно получить подставив в соответствующий список, полученный номер (3.18).

$$\begin{aligned}
Max[ListTA] \\
NB = Ordering[ListTA, -1] \\
ListW[1][[NB]] ListW[2][[NB]] ListW[3][[NB]] ListW[4][[NB]]
\end{aligned} \tag{3.18}$$

Полученные результаты на тренировочном множестве затем применяются на тестовом множестве, на котором значение показателя общей точности классификатора оказалась равна 65,2%.

В Приложении А.1 представлены первые 100 значений используемой базы кредитных историй из-за массивности данных. Обработав данные физических лиц по приведённой выше схеме, вычислена искомая оценка кредитоспособности. В Таблице 3.3 представлены оценки кредитоспособности для 100 клиентов банка по используемой базе кредитных историй (3.19).

$$Do[ListQtr[[NB, i]], \{i, 1, 100\}] \tag{3.19}$$

Значения кредитоспособности для 500 клиентов тестового множества представлены в Приложении А.5, которые вычислены с использованием следующего кода (3.20).

$$\begin{aligned}
 &ListQopttestset = Table[0, \{i, 1, 500\}]; \\
 &Do[ListQopttestset[[i]] = ListLevel2[[i, 1]] * ListW[1][[NB]] + \\
 &ListLevel2[[i, 2]] * ListW[2][[NB]] + \\
 &ListLevel2[[i, 3]] * ListW[3][[NB]] + \\
 &ListLevel2[[i, 4]] * ListW[4][[NB]], \{i, 1, 500\}
 \end{aligned} \tag{3.20}$$

Полученный класс кредитоспособности варьируемый по непрерывной шкале имеет большой потенциал для применения банками, в частности для:

1. дифференциации процентной ставки по кредиту;
2. визуализации допустимых вариантов кредитования для клиента;
3. управления атрибутами кредитного контракта на основе динамики кредитоспособности заёмщиков.

## 3.2 Система поддержки принятия решений по обоснованию процентной ставки физическим лицам

Существующие СППР по определению условий кредитования отсеивают некредитоспособных клиентов и устанавливают одинаковую процентную ставку для остальных, подающих заявку на один тип кредита. Рекомендации разработанной СППР выглядит следующим образом:

- заёмщики с кредитоспособностью ниже допустимого уровня не выдавать кредит;
- процентная ставка для остальных заёмщиков дифференцируется в соответствии с классом кредитоспособности.

Таблица 3.3: Класс кредитоспособности для первых 100 значений базы кредитных историй по модифицированному методу

№	Оценка	№	Оценка	№	Оценка	№	Оценка
1	0,56121	26	0,46758	51	0,59130	76	0,48683
2	0,54698	27	0,55339	52	0,45583	77	0,56648
3	0,58576	28	0,41885	53	0,60836	78	0,51708
4	0,47402	29	0,56910	54	0,48613	79	0,52297
5	0,56119	30	0,44996	55	0,72385	80	0,48617
6	0,49627	31	0,58669	56	0,41316	81	0,53885
7	0,47656	32	0,53590	57	0,51082	82	0,27683
8	0,59616	33	0,58648	58	0,39586	83	0,58702
9	0,59870	34	0,64429	59	0,50612	84	0,41589
10	0,68633	35	0,51335	60	0,34698	85	0,70700
11	0,58174	36	0,45401	61	0,54417	86	0,42291
12	0,55639	37	0,64129	62	0,48457	87	0,48774
13	0,52371	38	0,58733	63	0,42741	88	0,47819
14	0,62566	39	0,61436	64	0,44207	89	0,51217
15	0,50284	40	0,34830	65	0,56975	90	0,51138
16	0,63817	41	0,48383	66	0,56857	91	0,50114
17	0,41866	42	0,54530	67	0,54320	92	0,56644
18	0,45862	43	0,60059	68	0,51752	93	0,53983
19	0,56591	44	0,44675	69	0,61631	94	0,45357
20	0,55056	45	0,46438	70	0,55127	95	0,68530
21	0,50537	46	0,53182	71	0,47789	96	0,38644
22	0,52128	47	0,55417	72	0,53979	97	0,61102
23	0,47943	48	0,51350	73	0,58184	98	0,34948
24	0,58558	49	0,46711	74	0,48461	99	0,41518
25	0,59822	50	0,45971	75	0,50756	100	0,39729

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, кому следует отказать в выдаче кредита и значение процентной ставки по кредиту ( $P_j$ ) для каждого  $j$ -го кредитоспособного клиента.

Для определения зависимости процентной ставки от класса кредитоспособности, необходимо проанализировать затраты банка. В структуре затрат банка

от действий заёмщика зависят потери по ссудам. 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 потери по ссудам определяются, как функция  $(Q)$  (3.21).

$$c(Q) = \sum_{i=0}^k a_i * Q^i, \quad a_i \in R, \forall i = 0, 1, 2, \dots, k \quad (3.21)$$

Процентная ставка по кредиту может быть представлена, как сумма ожидаемых потерь по ссудам и остальных затрат по кредиту. В таком случае, 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 (3.22).

$$P_{min} = \frac{Sr_s + \sum_{i=1}^z d_i r_i + E}{M_{Total}}, \quad (3.22)$$

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.

В итоге процентная ставка для кредитоспособных клиентов определяется, как функция  $P(Q)$  (3.23).

$$P(Q) = P_{min} + \sum_{i=0}^k a_i * Q^i \quad (3.23)$$

Функцию процентной ставки можно составить следующим образом. На основании данных ЦБ РФ было получено значение минимальной процентной став-

ки в 13,01%. Статистика, на основе которой можно получить данную величину представлена в следующем параграфе. Средняя премия за риск составляет 13,7%. Область допустимых значений для процентной ставки варьирует от минимальной процентной ставки до суммы минимальной и удвоенного значения средней премии за риск, 40,41%. Используя значение кредитоспособности 0,4 как минимально допустимое, при котором 7% клиентов по тестовой выборке не получают кредит, а 0,8, как значение соответствующее минимальной процентной ставке в связи с тем, что ни один клиент из тестовой выборки не достиг данного значения, что может свидетельствовать о завышенности требований банка к идеальному заёмщику. Допустим, функция премии за риск для клиентов с кредитоспособностью выше допустимого уровня получилась линейной (3.24). Код для вычисления процентной ставки выглядит следующим образом (3.25).

$$P(Q) = 13,01\% + \frac{0,8 - Q}{0,2} * 13,7\% \quad (3.24)$$

$$P = Table[0, \{i, 1, 1000\}]; Do[P[j] = 13.01\% + (0.8 - Q[j]) * 5 * 13.7\%, j, 1, 1000] \quad (3.25)$$

Результаты системы поддержки принятия решений по определению условий кредита для первых 100 клиентов тестового множества представлены в Таблице 3.4. Стоит отметить, что на практике доля некредитоспособных заёмщиков, как правило, составляет менее 10% от числа поданных заявлений в банк [76]. Поэтому банк может использовать 10-ти процентную квантиль оценки кредитоспособности  $Q^{(10)}$  или в более общем случае  $x$ -процентную квантиль оценки кредитоспособности  $Q^{(x)}$ .

Для конкретного клиента функционирование разработанной СППР выглядит следующим образом. Клиент № 577 из вычислений, представленных в диссертации (клиент № 77 из тестовой выборки) подал заявку на кредит. На вход в систему поддержки принятия решений поступила рассматриваемая заявка. Некоторые из 16 используемых характеристик следующие: возраст ( $x_{13}^{577}$ ) – 27 лет; величина сбережений, в том числе и облигаций ( $x_6^{577}$ ) – до 50 тыс. рублей.

Таблица 3.4: Рекомендация системы поддержки принятия решений по выдаче кредита

№	решение	№	решение	№	решение	№	решение
1	29,37%	26	35,78%	51	27,31%	76	34,46%
2	30,34%	27	29,90%	52	36,59%	77	29,01%
3	27,69%	28	39,12%	53	26,14%	78	32,39%
4	35,34%	29	28,83%	54	34,51%	79	31,99%
5	29,37%	30	36,99%	55	18,23%	80	34,51%
6	33,82%	31	27,62%	56	39,51%	81	30,90%
7	35,17%	32	31,10%	57	32,82%	82	отказ
8	26,97%	33	27,64%	58	отказ	83	27,60%
9	26,80%	34	23,68%	59	33,14%	84	39,32%
10	20,80%	35	32,65%	60	отказ	85	19,38%
11	27,96%	36	36,71%	61	30,53%	86	38,84%
12	29,70%	37	23,88%	62	34,62%	87	34,40%
13	31,94%	38	27,58%	63	38,53%	88	35,05%
14	24,95%	39	25,73%	64	37,53%	89	32,73%
15	33,37%	40	отказ	65	28,78%	90	32,78%
16	24,10%	41	34,67%	66	28,86%	91	33,48%
17	39,13%	42	30,46%	67	30,60%	92	29,01%
18	36,39%	43	26,67%	68	32,36%	93	30,83%
19	29,05%	44	37,21%	69	25,59%	94	36,74%
20	30,10%	45	36,00%	70	30,05%	95	20,87%
21	33,19%	46	31,38%	71	35,07%	96	отказ
22	32,10%	47	29,85%	72	30,83%	97	25,96%
23	34,97%	48	32,63%	73	27,95%	98	отказ
24	27,70%	49	35,81%	74	34,61%	99	39,37%
25	26,83%	50	36,32%	75	33,04%	100	отказ

Нормировка характеристик осуществляется следующим образом. Для характеристики величина сбережений крайние категории, четвёртая и пятая получают оценку 0 и 1, соответственно. Остальные категории нормируются по среднему значению:

- категории 1 соответствует оценка  $(50000 - 0)/(2 * 500000) = 0,05$ ;
- категории 2 –  $(50000 + (250000 - 50000)/2)/500000 = 0,3$ ;
- категории 3 –  $(250000 + (500000 - 250000)/2)/(500000) = 0,75$ .

Для характеристики возраст, которая при увеличении до некоторого значения увеличивает значение класса  $Q$ , а после уменьшает, составлены нечёткие

множества молодой ( $Y$ ), среднего возраста ( $M$ ), преклонного возраста ( $O$ ) с функциями принадлежности, которые были приведены ранее (3.2-3.4).

Для вычисления результирующей функции принадлежности ( $\phi_{res}(y)$ ) сначала необходимо определить вывод по каждому из правил. Для первого правила  $\phi_{S^*}(y) = \text{Max}(\text{Min}(\phi_Y(x^*), \phi_S(y)), \text{Min}(\phi_O(x^*), \phi_S(y)))$ . Для второго правила  $\phi_{L^*}(y) = \text{Min}(\phi_M(x^*), \phi_L(y))$ . Тогда результирующая функция примет вид  $\phi_{res}(y) = \text{Max}(\phi_{S^*}(y), \phi_{L^*}(y))$ .

Для рассматриваемого случая нечёткие множества принимают следующие значения: (молодой)  $\phi_Y(x) = 0,65$ , (среднего возраста)  $\phi_M(x) = 0,35$ , (преклонного возраста)  $\phi_O(x) = 0$ . Функция вывода по первому правилу равна  $\phi_{S^*}(y) = \text{Max}(\text{Min}(0,65, \phi_S(y)), \text{Min}(0, \phi_S(y))) = \text{Min}(0,65, \phi_S(y))$ , по второму правилу  $\phi_{L^*}(y) = \text{Min}(0,35, \phi_L(y))$ . Результирующая функция принадлежности имеет следующий вид (3.26):

$$\phi_{res}(y) = \begin{cases} 0,65, & y \in [0, 0,35], \\ 1 - y, & y \in (0,35, 0,65), \\ 0,35, & y \in [0,65, 1] \end{cases} \quad (3.26)$$

Затем осуществляется дефазификация по методу центра тяжести, согласно которому момент результирующей функции принадлежности делится на площадь под графиком данной функции:

$$y_c = \frac{\int_0^1 y \phi_{res}(y) dy}{\int_0^1 \phi_{res}(y) dy},$$

где  $\phi_{res}(y)$  – результирующая функция принадлежности множеству  $y$ .

Стоит отметить, что итоговая оценка варьирует в диапазоне  $[\frac{1}{3}, \frac{2}{3}]$ , поэтому результат дефазификации необходимо нормировать в  $[0,1]$  с помощью следующей функции:  $f(y_c) = 3y_c - 1$ , где  $f : y_c \rightarrow [0,1]$ . Итоговое значение оценки характеристики возраст составило  $\frac{1}{3}$ . Данный результат в представленных ранее терминах разработанного метода может быть представлен следующим образом:  $f(y_c(y(x))) = q_{13}(x_{13}^{(577)}) = \frac{1}{3}$

Весовые коэффициенты для характеристик возраст и сбережения полученные путем настройки параметров по обучающей выборке равны 0,0245 и 0,03275,

соответственно. Получается, что вклад характеристик возраст и сбережения в результат классификации составил 0,00817 и 0,00164, соответственно. Итоговый класс с учётом остальных характеристик равен 0,56648.

На основе вычисленной ранее функции для процентной ставки по кредиту (3.24) искомая ставка для рассматриваемого клиента равна 29,01%. В случае, если клиент не предоставит информацию в банк о наличии сбережений, то процентная ставка по кредиту увеличится на 0,1%. В случае, если клиент предоставит информацию в банк о наличии сбережений в размере 200 тыс. рублей, то процентная ставка по кредиту уменьшится на 0,6%.

Так же предложен алгоритм, основанный 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. Помимо применения классификации физических лиц по кредитоспособности по непрерывной шкале отличие от существующих алгоритмов, где характеристики физического лица и атрибуты предполагаемого кредита анализируются вместе, состоит в осуществлении анализа кредитоспособности физического лица независимо от атрибутов предполагаемой кредитной сделки, что 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 степень рисковости 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 вычисляет диапазон варьирования атрибутов кредита 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  $T$  параллельно осям  $P$  и  $M$ . В результате СППР вычисляет диапазон варьирования для оставшихся атрибутов  $P \in [P'_{min}, P'_{max}]$ ,  $M \in [M'_{min}, M'_{max}]$ , которые не больше значений, полученных на предыдущем этапе ( $P_{min} \leq P'_{min}$ ,  $P'_{max} \leq P_{max}$ ,  $M_{min} \leq M'_{min}$ ,  $M'_{max} \leq M_{max}$ ).

**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 ( $M_j^{lim}$ ).

Стоит отметить, что после выдачи кредита финансовое положение заёмщика может измениться, что требует соответствующего изменения в оценке кредито-

способности [89]. Поэтому 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 новые значения процентной ставки по кредиту ( $P_j(t) = F_P^{-1}(Q_j(t), T_j(t), M_j(t))$ ,  $T_j(t)$  – срок кредита для  $j$ -го клиента на момент  $t$ ) и 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.

В предложенной модели СППР параметры рисковости кредитного контракта и характеристики клиента раздельно анализируются, чтобы получить множество поверхностей уровня  $F(P, T, M) = Q$ , где  $T$  – продолжительность кредита,  $M$  – сумма кредита,  $P$  – процентная ставка. Данные поверхности уровня отображают одинаково оплаченный риск при фиксированном уровне кредитоспособности (3.27).

$$F(P, T, M) = Q \quad (3.27)$$

При подстановке конкретного значения кредитоспособности клиента  $Q_i$  получается спектр условий кредитования, соответствующий индивидуальному уровню кредитоспособности, при котором риск банка одинаково оплачен (3.28).

$$F(P, T, M) = Q_i \quad (3.28)$$

На Рисунке 3.2 изображены линии уровня, выражающие зависимость процентной ставки от суммы кредита при различной кредитоспособности (3.29). Наклонная линия показывает зависимость кредитного лимита от кредитоспособности и процентной ставки (3.30). Клиент может выбрать любую комбинацию условий кредитования на соответствующей линии уровня до кредитного лимита, например комбинацию  $(M_1, P_1)$  (Рис. 3.2).

$$Q_1 < Q_2 < Q_3 \quad (3.29)$$

$$M_1^{lim} < M_2^{lim} < M_3^{lim} \quad (3.30)$$

Кредитный риск, как правило, является возрастающей функцией от суммы кредита и от срока кредита. Однако если проанализировать зависимость процентной ставки от объема кредита за период с 01 июля по 30 сентября 2015 года, то можно увидеть, что данная зависимость не всегда является монотонно возрастающей.

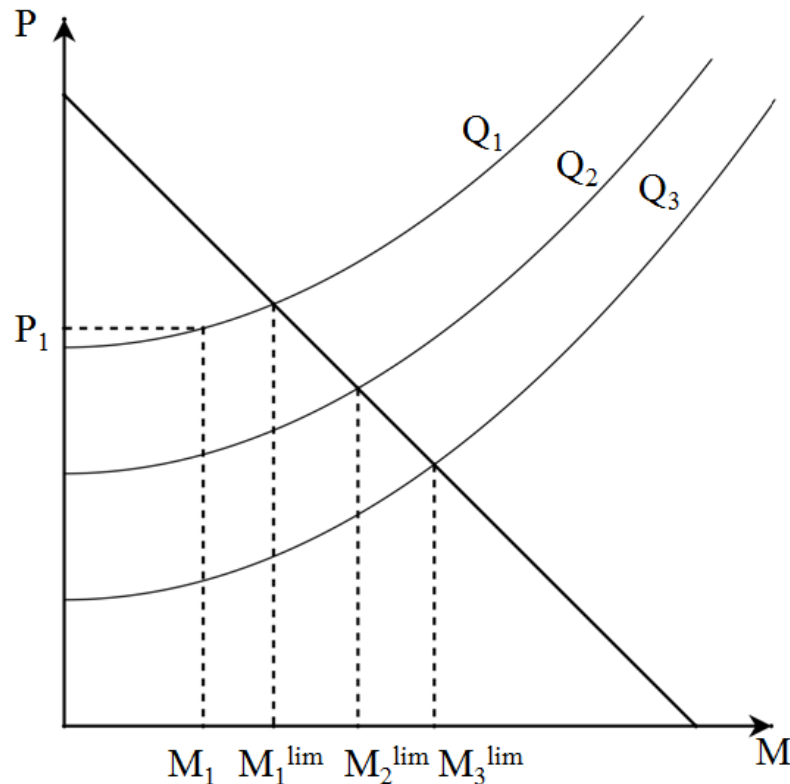


Рисунок 3.2: Линия уровня процентной ставки по кредиту от суммы кредита при различном значении кредитоспособности

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  $(P, T, M)$  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 (3.31):

$$P_i = F_P^{-1}(Q, T, M) = \sum_{i=0}^n (a_i M^i + b_i T^i + d_i Q^i),$$

$$a_i, b_i, d_i \in R, i = 0, 1, \dots, n, \quad (3.31)$$

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

– для вычисления кредитного лимита (3.32):

$$M_i^{lim} = F_M^{-1}(Q, T, P) = \sum_{i=0}^n (w_i P^i + u_i T^i + v_i Q^i),$$

$$w_i, u_i, v_i \in R, i = 0, 1, \dots, n, \quad (3.32)$$

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

Ожидается, что чем больше сумма кредита, тем выше потери по ссудам. Аналогичная зависимость ожидается от срока кредитования. Однако, чем выше кредитоспособность, тем ниже ожидаемые потери по кредиту.

Управлять атрибутами кредита на основе динамики класса кредитоспособности заёмщиков банк может следующим образом. Допустим, класс кредитоспособности  $j$ -го заёмщика за период  $\Delta t$  изменилась с  $\alpha$  до  $\beta$ , 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  $\Delta M$  (3.33):

$$\begin{aligned} \Delta M_j^{lim}(t) &= M_j^{lim}(t) - M_j^{lim}(t-1) = \\ &= F_M^{-1}(P(t), T(t), \beta) - F_M^{-1}(P(t-1), T(t-1), \alpha), \end{aligned} \quad (3.33)$$

в случае если прочие параметры неизменны данное выражение можно привести к следующему виду (3.34):

$$\begin{aligned}\Delta M_j^{lim}(t) &= \sum_{i=0}^n (w_i P^i + u_i T^i + v_i \beta^i) - \sum_{i=0}^n (w_i P^i + u_i T^i + v_i \alpha^i) = \\ &= \sum_{i=0}^n v_i (\beta^i - \alpha^i)\end{aligned}\quad (3.34)$$

2. изменить процентную ставку по кредиту на величину  $\Delta P_j(t)$  (3.35):

$$\begin{aligned}\Delta P_j(t) &= P_j(t) - P_j(t-1) = \\ &= F_P^{-1}(M(t), T(t), \beta) - F_P^{-1}(M(t-1), T(t-1), \alpha),\end{aligned}\quad (3.35)$$

при конкретном значении суммы кредита ( $M$ ) величина  $\Delta P_j(t)$  может выступать, как разность процентных ставок по кредиту при величине кредита  $M$  и соответствующих классах кредитоспособности (3.36):

$$\begin{aligned}\Delta P_j(t) &= \sum_{i=0}^k (q_i M^i + p_i T^i + h_i \beta^i) - \sum_{i=0}^k (q_i M^i + p_i T^i + h_i \alpha^i) = \\ &= \sum_{i=0}^k h_i (\beta^i - \alpha^i)\end{aligned}\quad (3.36)$$

Изменение уровня кредитоспособности можно представить, как сдвиг поверхности уровня по осям, которые характеризуют условия кредитования. Например, в осях  $(T, M)$  при фиксированной процентной ставке клиентам с различным уровнем кредитоспособности ( $Q_1 < Q_2 < Q_3 < Q_4 < Q_5$ ) будут соответствовать линии уровня условий кредитования, как представлено на Рисунке 3.3.

Функционирование предложенной системы поддержки принятия решений по предложенному алгоритму продемонстрировано на следующем примере. Допустим, что банк вычислил линий уровня, которые отображают соотношение процентной ставки по кредиту и суммы кредита, которые по премии за риск соответствуют классу кредитоспособности клиента  $F(P, M) = Q$ . На Рисунке 3.4

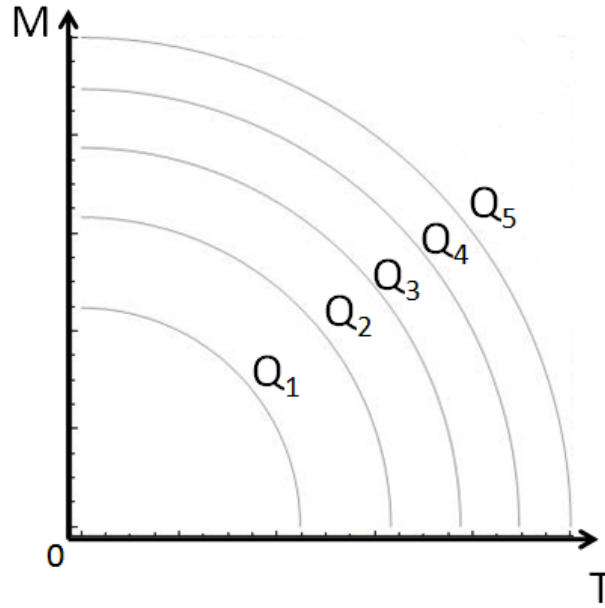


Рисунок 3.3: Линии уровня условий кредитования при фиксированной процентной ставке

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 лимите по кредитной линии 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  $M_3^{lim} = F_M^{-1}(P_3, Q_j(t_3))$ . Заём-

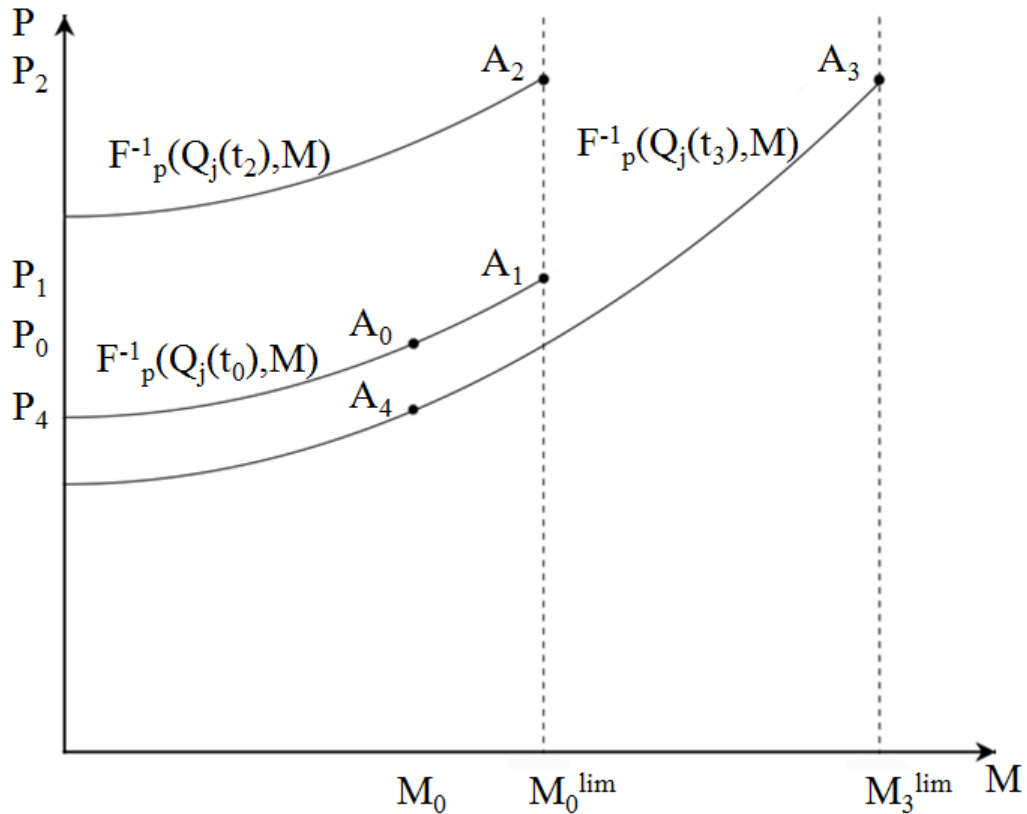


Рисунок 3.4: Комбинации условий кредитования

цику 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  $P_4 = F_p^{-1}(Q_j(t_4), M_0)$ .

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. Клиент вы-

берет 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 возможность подобного выбора может выступать в качестве фактора для клиента при поиске кредитора.

Применить разработанный метод к портфелям однородных ссуд физических лиц по классификации Центрального Банка Российской Федерации [1] можно следующим образом. Жилищные и ипотечные ссуды связаны с большими суммами кредита. Сбои и ошибки в работе автоматизированной системы по выдаче кредита при данном типе кредитования обладают высокой стоимостью для банка. Автокредиты в среднем связаны с меньшей суммой займа, однако при отсутствии обеспечения стоимость ошибки при полностью автоматизированной системе по выдаче кредита, также очень высокая. В перечисленных ссудах целесообразно применять метод оценки кредитоспособности по непрерывной шкале в системе поддержки принятия решений по обработке заявлений на кредит, где окончательное решение по выдаче кредита осуществляется экспертом банка. Учитывая, что при данных типах кредитования задействованы большие суммы, система мониторинга динамики кредитоспособности является востребованной. Более того, финансовая нагрузка при рассматриваемых типах кредитования высокая, поэтому визуализация возможных условий кредитования будет востребована среди заёмщиков для выбора оптимального варианта кредитования. Удобство выплаты кредита заёмщиком повышает вероятность своевременных выплат по кредиту.

Наибольшую нагрузку при формировании резерва на возможные потери по ссудам физических лиц на 1.07.15 для банков оказывают прочие потребительские ссуды - 86,9%. Управление рисками в данной категории является наиболее востребованным для банков. Все 3 вышеперечисленных варианта применения метода оценки кредитоспособности по непрерывной шкале востребованы в данной категории кредитования.

По перечисленным вариантам можно сделать вывод, что применение оценки кредитоспособности по непрерывной шкале существенно увеличивает возможности банка по управлению кредитным риском. Приведенная ранее схема автоматизированного установления процентной ставки (Рис. 1.1) может быть моди-



фицирована с учётом возможностей, которые предоставляет оценка по непрерывной шкале в виде приведённых выше моделей (Рис. 3.5) Применение оценки кредитоспособности для дифференциации процентной ставки по предложенной ранее ценовой стратегии стимулирует клиентов предоставлять информацию в банк, чтобы получить более дешёвый кредит. Данный эффект позволит снизить затраты банка на сбор информации.

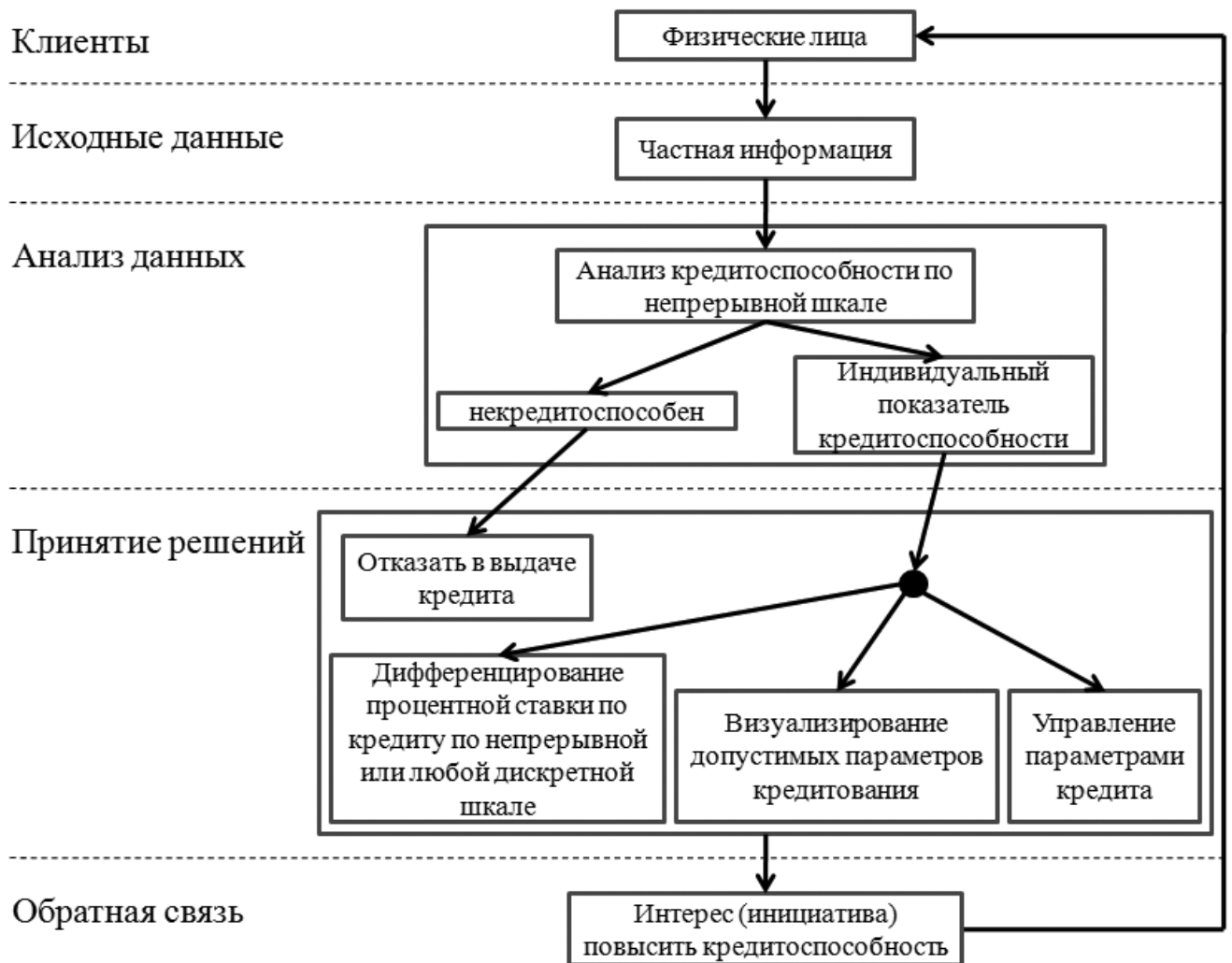


Рисунок 3.5: Схема автоматизированного взаимодействия с физическими лицами в кредитовании при использовании оценки кредитоспособности по непрерывной шкале

Таким образом, банкам для организации кредитования на индивидуальных условиях необходимо использовать методы классификации физических лиц по кредитоспособности по непрерывной шкале, которые необходимо реализовать в системе поддержки принятия решений. Применение подобной СППР позволит каждому клиенту на основе набора индивидуальных характеристик (частной информации) определяется индивидуальный класс кредитоспособности, по ко-

торому вычисляется премия за кредитный риск и индивидуальная процентная ставка по кредиту. Таким образом, при схожих затратах банка на сбор и анализ частной информации физических лиц банк сможет получить более информативную оценку кредитоспособности, что способствует более эффективному управлению кредитным риском и обеспечивает более стабильное развитие экономики.

Учитывая высокую значимость сектора потребительского кредитования для экономики дальнейшая разработка предложенного метода классификации физических лиц по кредитоспособности и решение смежных вопросов являются перспективными направлениями исследования. В качестве одного из подобных вопросов может быть следующий: внедрение новой системы выдачи кредитов физическим лицам потребует от коммерческих банков изменения архитектуры управления бизнесом [10]. Для дальнейшего развития метода классификации физических лиц по кредитоспособности по непрерывной шкале целесообразно использовать нечёткую логику.

### 3.3 Экономический эффект и эффективность

В общем, точность методов классификации объектов оценивается на основе вероятности верно классифицированных клиентов, он же показатель общей точности (total accuracy). Ошибки при бинарной классификации делятся на ошибки первого и второго рода. Более наглядно можно записать варианты прогноза и фактическую реализацию процесса в виде Таблицы 3.5, по которой вычисляется доля истинно положительных примеров (3.37), которая является чувствительностью модели, долю ложно положительных примеров (3.38) и специфичность модели (3.39), на основе которых строится ROC-кривая. Данная кривая так же используется при определении эффективности классификатора.

Таблица 3.5: Истинность классификации

Прогноз	Реализация	
	Положительно	Отрицательно
Положительно	TP	FP
Отрицательно	FN	TN

где TP – положительные наблюдения, классифицированные как положительные.

TN – отрицательные наблюдения, классифицированные как отрицательные.

FN – положительные наблюдения, классифицированные как отрицательные (ошибка первого рода).

FP – отрицательные наблюдения, классифицированные как положительные (ошибка второго рода).

$$TPR = \frac{TP}{TP + FN} \quad (3.37)$$

$$FPR = \frac{FP}{TN + FP} \quad (3.38)$$

$$S_p = \frac{TN}{TN + FP} \quad (3.39)$$

При сравнении классификаторов ROC-кривой может быть недостаточно. В случае, если ROC-кривая для первого классификатора превосходит во всех точках значения для второго, то первый классификатор лучше (Рис. 3.6). Однако в случае пересечения графиков необходимы дополнительные инструменты (Рис. 3.7). Для того, чтобы сравнивать классификаторы используют площадь под ROC-кривой (Area Under ROC curve).

При сравнении классификаторов так же используют показатель среднегеометрического значения между долей истинно положительных примеров и чувствительностью модели (3.40) [136]. Однако данный показатель не учитывает, что затраты для разных типов ошибок различны.

$$Gmean = \sqrt{\frac{TP}{TP + FN} \frac{TN}{TN + FP}} \quad (3.40)$$

Общая точность классификатора (total accuracy) вычисляется, как отношение числа верно предсказанных наблюдений и общего числа наблюдений (N) (3.41) [57].

$$TA = \frac{TP + TN}{N} \quad (3.41)$$

При сравнении методов классификации апробированных на используемой базе кредитных историй, по показателю общей точности, составлена Таблица 3.6 (с использованием источника [45]).

Таблица 3.6: Точность методов классификации

Название метода	Точность (Total accuracy)
К ближайших соседей	62,0%
Опорные векторы	63,5%
Модиф. рандомизированные показатели	65,2%

В сравнении с методом опорных векторов и К ближайших соседей, разработанный модифицированный метод рандомизированных показателей показал более высокое значение показателя общей точности на тестовом множестве.

Представим потери по ссудам, как функцию от показателя общей точности классификатора (3.42):

$$C(TA) = (1 - TA) * M, \quad (3.42)$$

где  $TA$  – показатель общей точности классификатора,  $M$  – сумма выданных кредитов.

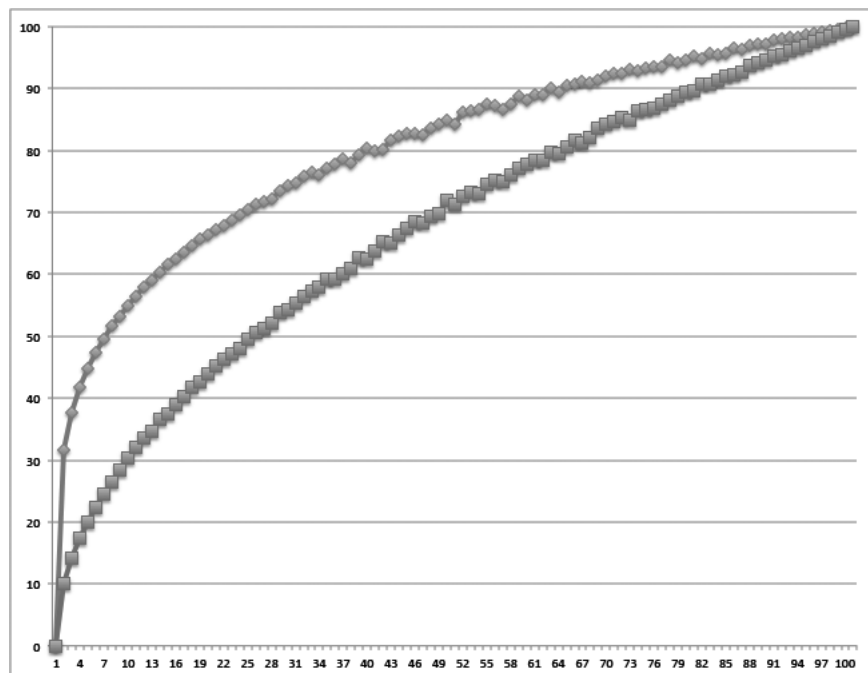


Рисунок 3.6: Две непересекающиеся ROC-кривые (%)

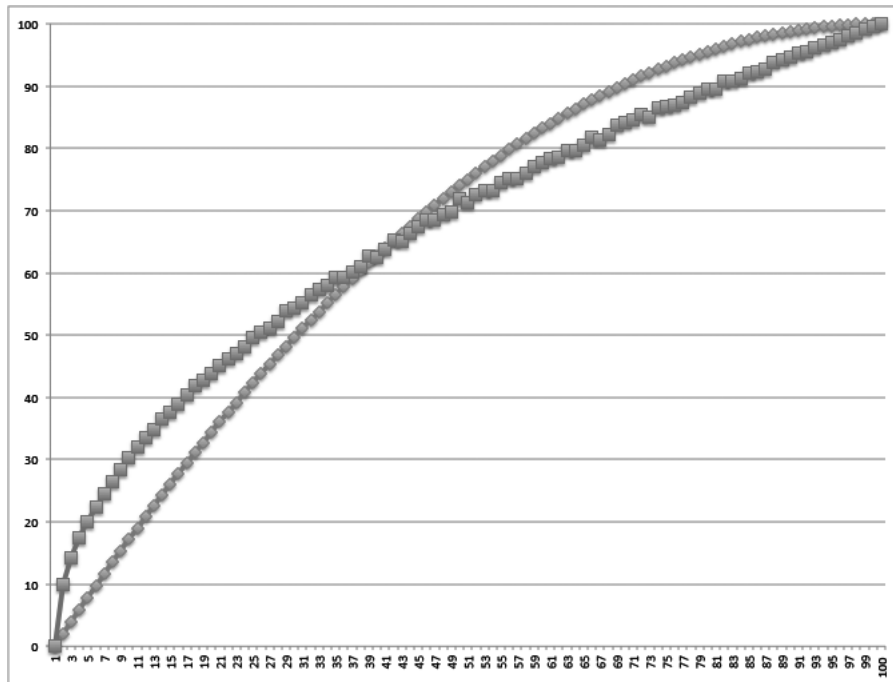


Рисунок 3.7: Две пересекающиеся ROC-кривые (%)

Тогда снижение потерь по ссудам при увеличении показателя общей точности от  $TA_1$  до  $TA_2$  с прежней суммой выданных кредитов может быть записано следующим образом:

$$\begin{aligned} \Delta C &= \frac{C(TA_1) - C(TA_2)}{C(TA_1)} = \frac{(1 - TA_1) * M - (1 - TA_2) * M}{(1 - TA_1) * M} = \\ &= 1 - \frac{1 - TA_2}{1 - TA_1} \end{aligned} \quad (3.43)$$

Переход к предложенному методу от метода опорных векторов без изменения политики банка в части дифференцирования процентной ставки по кредиту обеспечит снижение потерь по ссудам на 4,66% (согласно формуле 3.43).

Эффективность от внедрения предложенного метода можно оценить посредством анализа статистических данных ЦБ РФ. Вычислим средневзвешенную стоимость капитала (WACC). Для этого определим структуру и стоимость капитала банков. Среднее значение норматива достаточности собственного капитала ЦБ РФ (Н1) на 01.01.2015 составило 12,5% [21]. Стоимость собственного капитала можно определить как сумму безрисковой ставки, премии за риск и требуемой нормы доходности. В качестве безрисковой ставки для коммерческого банка может выступать ключевая ставка ЦБ величиной в 11% годовых. В

условиях нестабильности экономики банковская деятельность является высокорисковой. Требуемая доходность, которую инвесторы предъявляют к банковскому капиталу, как правило, существенно выше, нежели в других областях. Суммарно премию за риск без учёта потерь по ссудам и требуемую прибыль оценим на уровне 9% годовых. В таком случае стоимость (требуемая норма доходности) собственного капитала составит 20% годовых. Обеспеченность кредитов депозитами на 01.01.2015 составила 80,9% [21]. Среднегодовая процентная ставка по депозитам на срок до 1 года (включая до востребования) за 2015 составила 9,27%<sup>4</sup>. Допустим, что оставшийся капитал банк привлекает под те же 9,27%. В таком случае средневзвешенная стоимость капитала составит 10,61% годовых.

$$c = 12,5\% * 20\% + 87,5\% * 9,27\% = 10,61\% \quad (3.44)$$

Стоит учитывать, что не все средства банк может использовать для выдачи кредитов. Обязательные резервные требования составляют 4,25%. В добавок резервы на возможные потери по ссудам на 01.01.2015 составили 6,5% [1]. За 2015 год средняя процентная ставка по кредиту составила 26,71%. Допустим, что затраты на ведение банковской деятельности составляют 1% от объема располагаемых банком средств (обозначим, как  $z$ ). Уровень кредитного риска можно вычислить по формуле (3.45). В итоге средняя премия за кредитный риск примерно равна 13,7% (3.45). Относительное снижение уровня кредитного риска при переходе к разработанному методу от метода опорных векторов составит 0,638% ( $13,7\% * 4,66\%$ ).

$$r = 26,71\% - \frac{c + z}{1 - res} \quad (3.45)$$

где  $res$  – суммарные резервы, в текущем примере 10,75%.

В последнее время исследователи при классификации физических лиц по кредитоспособности активно развивают метод опорных векторов. Допустим, что в секторе необеспеченных потребительских кредитов, который в рассматриваемый год составил 6,4 трлн. рублей, все заявки на кредит анализируются с помощью метода опорных векторов. Стоит отметить, что методика анализа

---

<sup>4</sup>Вычислено по данным ЦБ РФ. Режим доступа: [http://www.cbr.ru/statistics/?PrtId=int\\_rat](http://www.cbr.ru/statistics/?PrtId=int_rat)

частной информации физических лиц конкретного банка является коммерческой тайной. Экономический эффект от повсеместного использования предложенного метода в рассматриваемом секторе при прежнем (усредненном) ценообразовании кредита и прежнем объеме сектора составит 40,86 млрд. рублей экономии на потерях по ссудам.

Внедрение предложенной системы поддержки принятия решений может осуществляться полностью в рамках банка, а так же с помощью сторонних организаций. Оценить рентабельность вложений во внедрение в банке полученных результатов без участия сторонних компаний можно осуществить следующим образом. Предположим, что система поддержки принятия решений по определению процентной ставки разрабатывается год, а затем используется три года, после чего требует обновления, которое так же осуществляется год.

Затраты на обновление действующей системы поддержки принятия решений по определению процентной ставки следующие:

- заработная плата работы восьми программистов – 150 тыс. рублей в месяц каждому;
- арендная плата 80 кв. м. офисных помещений – 1 тыс. рублей за кв. м. в месяц;
- арендная плата необходимой техники, а именно компьютера, монитора и т.д. – 2 тыс. руб. в мес. на сотрудника;
- оплата программного обеспечения – 800 тыс. рублей.

Итого:  $8 \cdot 12 \cdot 150$  тыс. руб. +  $80 \cdot 12 \cdot 1$  тыс. руб. +  $8 \cdot 12 \cdot 2$  тыс. руб. + 800 тыс. руб. = 16, 352 млн. руб.

Допустим, что на разработку новой системы поддержки принятия решений банку потребуется на два программиста больше, что увеличивает затраты на 25%. Таким образом банку для разработки и внедрения предлагаемой в работе системы поддержки принятия решений требуется дополнительно инвестировать 4,088 млн. руб. по сравнению с плановым обновлением действующей системы.

Денежный поток во времени распределен следующим образом:

- в нулевой момент осуществляются затраты для разработки и внедрения СППР 4,088 млн. рублей;

- в конце первого года новая СППР начинает функционировать;
- в конце второго года банк получает приток денежных средств  $\Delta C * M$  за счет снижения потерь по ссудам;
- в конце третьего года банк получает приток денежных средств  $\Delta C * M$ ;
- в конце четвертого года банк получает приток денежных средств  $\Delta C * M$ , после чего СППР требует обновления.

На основе ранее представленной статистики вычислена требуемая норма доходности, равная 33,7% (20% + 13,7%). В таком случае окупаемость инвестиции в предлагаемую систему поддержки решений достигается при портфеле потребительских кредитов банка в размере 496,4 млн. рублей на протяжении трёх лет при изначальном использовании метода опорных векторов или за один год при размере портфеля потребительских кредитов в 1,15 млрд. рублей. При размере портфеля потребительских кредитов в 700 млн. рублей в среднем за год на протяжении трех лет эффективность от вложенных инвестиций составит примерно 41%.

Стоит отметить, что вероятность верно классифицированных наблюдений не отражает всех преимуществ метода классификации по непрерывной шкале. Представленный факт связан с тем, что бинарная классификация является лишь одним из вариантов использования данного метода. Однако при увеличении числа классов вероятность попадания в конкретный класс снижается. Поэтому вероятность верно предсказанных исходов не корректно использовать при сравнении эффективности методов классификации с различным числом градаций, а так же при сравнении методов классификации по непрерывной шкале, что показано на следующих вычислениях.

Зададим вероятностное пространство  $(\Omega, \mathcal{A}, P)$ :

- $\Omega = \{A_1, A_2, \dots, A_n\}$  – множество элементарных исходов, классов, например степень кредитоспособности;
- $\mathcal{A}$  – сигма-алгебра подмножеств  $\Omega$ ;
- $P(\xi = A_i)$  – вероятность того, что клиент принадлежит классу  $A_i$ .



Рассмотрим случай дискретного равномерного распределения вероятности  $P(\xi = A_i) = \frac{1}{n}, \forall i = 1, 2, \dots, n$ . При бинарной классификации ( $n = 2$ ) вероятность принадлежности конкретному классу равна 50%:  $P(A_j) = \frac{1}{2}$ . При наличии трёх классов ( $n = 3$ ) данная вероятность равна 33,3%:  $P(A_j) = \frac{1}{3}$ . При наличии пяти классов данная вероятность равна 20%:  $P(A_j) = \frac{1}{5}$ . При увеличении числа классов данная вероятность стремится к нулю (3.46).

$$\lim_{n \rightarrow +\infty} P(\xi = A_i) = \lim_{n \rightarrow +\infty} \frac{1}{n} = 0 \quad (3.46)$$

Получается, что при использовании вероятности верно классифицированных клиентов, в качестве основы показателя точности, то чем больше классов, тем в среднем меньше точность метода.

Данный результат может быть получен иным образом. Допустим, что  $\xi$  – непрерывная случайная величина. Вероятность того, что  $\xi$  попадёт в отрезок от  $\alpha$  до  $\beta$  вычисляется по формуле (3.47).

$$P(\alpha \leq \xi \leq \beta) = \int_{\alpha}^{\beta} f(x) dx \quad (3.47)$$

где  $f(x)$  – плотность распределения случайной величины. Формулу можно записать с помощью функции распределения  $F_{\xi}(x)$  (3.48).

$$P(\alpha \leq \xi \leq \beta) = F_{\xi}(\beta) - F_{\xi}(\alpha) = \beta - \alpha. \quad (3.48)$$

В случае, если отрезок состоит из одной точки ( $\alpha = \beta = x_0$ ), вероятность попадания в данный отрезок равна нулю:  $P(\xi = x_0) = 0$ .

Таким образом, вероятность верно классифицированных клиентов некорректно использовать, в качестве единственной меры точности методов классификации с различными шкалами. При увеличении числа градаций вероятность попадания в конкретную категорию снижается. В случае сравнения методов классификации с непрерывной шкалой, вероятность верно классифицированных клиентов невозможно использовать, так как априорная вероятность соответствия конкретному классу равна нулю. В связи с этим далее представлен другой ожидаемый положительный эффект от применения разработанного метода при принятии решения о выдаче кредита.

Другим положительным эффектом от применения разработанного метода является увеличение финансовой стабильности банка в секторе кредитования физических лиц, что актуально, потому что за 2014 и 2015 годы более чем в 3,5 раза увеличилось число отозванных лицензий (не считая отзыв по причине реорганизации и ликвидации юридического лица) по сравнению с 2012 и 2013 годами.

Разработанный метод лежит в основе СППР по определению процентной ставки. Сравним существующие СППР и разработанную СППР составив пример однопериодного кредитования физических лиц для двух банков, смоделировав риск с помощью метода Монте-Карло.

Смоделируем с помощью метода Монте-Карло однопериодную модель кредитования физических лиц для двух банков, использующих существующую СППР по определению условий кредита и разработанную СППР, соответственно.

Каждый банк в начале периода берет деньги на депозит под некоторую процентную ставку и выдаёт их в форме кредитов. В конце периода банк собирает процентные платежи по кредиту и расплачивается по депозитам. Пусть  $u$  - случайная величина, характеризующая среднюю долю потерь по ссудам на единицу выданных средств в банке с распределением  $N(\phi, \sigma^2) * I[0,1]$ , где  $\phi$  - средняя величина премии за риск по сектору кредитования физических лиц. Разность средней доли потерь по ссудам и средней премии за риск равно  $d = u - \phi$ . Случайная величина  $d$  имеет распределение  $N(0, \sigma^2) * I[-\phi, 1 - \phi]$ .

Допустим, банк действует в условиях составленных ранее на основе данных ЦБ РФ:

- собственный капитала банка составляет 12,5% от величины активов;
- депозиты составляют 87,5%;
- процентная ставка по депозитам равна 9,27%;
- резервы составляют 10,75%, то есть только 89,25% средств банк выдаёт в виде кредитов;
- процентная ставка по кредиту составляет 26,71%;

– средняя премия за кредитный риск равна 13,7%.

В конце периода банк получает  $(26,71\%-c)*89,25\%$  или, что тоже самое,  $(13,01\%-d)*89,25\%$  со всего объема активов и должен отдать по депозитам  $9,27\%*87,5\%$ . В случае, если разность данных величин равна 0, то банк останется без прибыли и убытков, то есть величина активов останется неизменной. Если собственные средства банка будут ниже минимального допустимого значения достаточности собственных средств, то у него отзовут лицензию. Минимальное допустимое значение достаточности собственных средств (Н1.0) в 2015 году составляло 10% [21]. Таким образом, если разность данных величин будет меньше минус 2,5%, что достигается при  $d > 7,4\%$ , то у банка отзовут лицензию в виду нарушения норматива достаточности собственных средств (Н1.0).

Процентная ставка по кредиту и премия за риск определяются в банках следующим образом:

1. В первом случае банк применяет усреднённое ценообразование кредита физическим лицам, подающим заявку на один тип кредита, что соответствует использованию систем поддержки принятия решений по определению условий кредитования, основанных на существующих моделях классификации физических лиц по степени кредитоспособности. Процентная ставка равна средней по сектору кредитования физических лиц. Премия за риск, так же равна средней и обозначается  $\phi$ . Взяв стандартное отклонение равное 8% ( $\sigma = 8\%$ ) и осуществив 10000 повторов вероятность, что банк не сможет удовлетворять требованиям ЦБ достаточности собственных средств (Н1.0) составила 17,5%.
2. Во втором случае банк применяет дифференцированное ценообразование кредита физическим лицам в зависимости от уровня кредитоспособности, что соответствует применению разработанной системы поддержки принятия решений. Средняя премия за риск в банке равна  $\phi'$ . Очевидно, что при возможности дифференцировать процентную ставку стандартное отклонение разности средней процентной ставки по банку и средней доли потерь по ссудам можно снизить вплоть до значений близких к нулю. Допустим, что подобное дифференцирование снизило разброс разности средней процентной ставки по банку и средней доли потерь по ссудам

всего лишь на 1%. По результатам моделирования вероятность, что банк не сможет удовлетворять требованию центрального банка по достаточности собственных средств (Н1.0) составила 14,61%.

Таким образом при переходе от существующих СППР к разработанной вероятностной, что у банка, действующего в секторе потребительского кредитования, отзовут лицензию из-за не соответствия требованию центрального банка по достаточности собственных средств Н1.0 снизится на 2,89%. Если инвестировать средства в развитие предложенной СППР, снижая стандартное отклонение разброса разности средней процентной ставки по банку и средней доли потерь по ссудам, то подобный банк будет более стабилен к колебаниям рыночной конъюнктуры. Например, при снижении стандартного отклонения на 3% в абсолютном выражении, менее чем на 40% относительно исходной величины, вероятность отзыва лицензии снижается почти в 3 раза до значения 6,57%, то есть снижение составляет 10,93%.

Стоит отметить, что с первого января 2016 года минимальное значение норматива достаточности собственных средств снизилось до 8%, свидетельствуя о том, что банки не справляются с выполнением предыдущего норматива в 10%. Данный факт увеличивает актуальность внедрения банками разработанного метода и системы поддержки принятия решений.

Помимо коммерческих банков применение результатов работы актуально для ЦБ РФ для мониторинга уровня кредитного риска портфеля кредитов, выданных физическим лицам.

### **3.4 Особенности ценообразования кредита для физических лиц**

В текущем разделе предстоит определить особенности ценообразования кредита при использовании разработанного метода и системы поддержки принятия решений. Особое внимание уделено вопросу уменьшения негативных эффектов асимметрии информации.

В теории, согласно гипотезе об эффективном рынке капитала, цена актива должна отражать всю информацию в любое время [68]. Наиболее распростра-

нённая ценовая стратегия на рынке потребительского кредитования, которая заключается в установлении единой процентной ставки (“house rate”) [88], далека от соответствия данной гипотезе. Позднее, в работе Гроссмана и Стиглица о невозможности информационно эффективных рынков [75], было показано, что для установления эффективных рынков необходимо, чтобы затраты на получение информации были равны 0. После опубликования указанной работы приведённая формулировка гипотезы эффективного рынка стала называться строгой версией [69]. В текущей работе ориентиром при построении ценовой стратегии по выдаче кредита физическим лицам является указанная гипотеза, то есть отражение всей доступной информации о клиенте в процентной ставке по кредиту. Более того, в последнее время концепция доверия стала весьма популярна в информационных системах [102], в частности, в мобильных банковских системах [36, 93, 105, 107] и методах классификации физических лиц по кредитоспособности [32]. Подобная концепция имеет потенциал для применения при ценообразовании кредита в потребительском кредитовании, обеспечивая снижение неблагоприятного отбора и стимулирования клиентов предоставлять частную информацию в банк.

Банки расходуют средства для сбора информации о клиентах. Однако в большинстве банков данная информация не используется для дифференциации процентной ставки по кредиту для физических лиц (за исключением выдачи крупных сумм, как в ипотечном кредитовании). В случае, если клиент самостоятельно предоставляет в банк частную информацию, то у банка отсутствует потребность расходовать средства на сбор данной информации.

В работе [54] на основе данных испанских фирм получено, что долгосрочные кредитные отношения (relationship lending) снижают процентные ставки по кредиту. Кроме того, в секторе корпоративного кредитования связи между банком и заёмщиком снижают процентные ставки по кредиту [67]. В связи с этим, банку целесообразно формировать доверительные отношения с клиентами для стимулирования последних предоставлять частную информацию, обеспечивая снижение процентной ставки по кредиту.

Особое значение имеет рассмотрение вопроса ценообразования кредита именно для физических лиц в связи со следующим:

1. для фирм существуют различные формы отчётности, такие как бухгалтерская, налоговая и т. п., которые необходимо соблюдать, что делает фирмы информационно более прозрачными, нежели физических лиц;
2. средний размер кредита фирмам существенно выше, нежели физическим лицам, поэтому банку необходимо больше ресурсов израсходовать на единицу выданных кредитных средств для принятия обоснованных решений при выдаче кредита последним;
3. ценообразование кредита корпоративным клиентам в большинстве случаев состоит в дифференциации процентной ставки по кредиту в соответствии с кредитным риском (*risk based pricing*). Однако при кредитовании физических лиц подобная практика малоприменима.

Как было упомянуто ранее наиболее часто встречающееся ценообразование кредита состоит в отсеке наименее кредитоспособных клиентов и выдаче остальным кредита по единой процентной ставке. При подобном ценообразовании кредита часть заёмщиков переплачивает за кредит, а часть недоплачивает. Справедливо по отношению к заёмщикам устанавливать процентную ставку в соответствии с себестоимостью кредита, которая зависит от потерь по ссудам. Однако фактические потери по ссудам известны после окончательного исполнения контракта по кредиту. На момент выдачи кредита банк может оценивать потери по ссудам, классификации физических лиц по кредитоспособности, которая осуществляется по располагаемой информации о клиентах. Таким образом классификации физических лиц по кредитоспособности может быть использована банком при формировании материального стимула клиентам предоставлять частную информацию.

Преградой для заёмщика предоставлять частную информацию при дифференциации процентной ставки по кредиту является опасение, что данная информация понизит класс кредитоспособности и увеличит процентную ставку. Предлагаемая ценовая стратегия по выдаче кредита сформулирована состоит в следующем. Отсутствие информации определяет наихудший класс кредитоспособности, что соответствует наихудшим условиям кредитования. Предоставляемая частная информация не может сделать процентную ставку по кредиту выше (3.49). В случае, если предоставленная информация увеличивает класс

кредитоспособности клиента, то процентная ставка по кредиту становится ниже (3.50).

$$P(I_b \cup I_s) \leq P(I_b) \quad (3.49)$$

$$\widehat{Q}(I_b \cup I_s) > \widehat{Q}(I_b), P(\widehat{Q}(I_b \cup I_s)) < P(\widehat{Q}(I_b)), I_s = \{I_p, I_n\}, I_p \not\subseteq I_b \quad (3.50)$$

Клиент банка, в котором ценовая стратегия основана на данных принципах, материально заинтересован предоставлять частную информацию, потому что это может только понизить процентную ставку по кредиту при предоставлении информации, которая увеличивает класс кредитоспособности и оставляет неизменной при предоставлении негативной информации. В данных условиях, клиент уверен, что предоставляемая частная информация не сделает условия кредитования менее выгодными. В результате для клиента более чётко формируется инициатива предоставлять частную информацию в банк.

Согласно сформулированной ценовой стратегии банк предоставляет физическим лицам “премии” за предоставленную информацию, дифференцируя процентную ставку по кредиту в соответствии с классом кредитоспособности, чтобы не расходовать средства на сбор частной информации клиентов в секторе потребительского кредитования. Реализация предложенной ценовой стратегии зависит от возможностей банка отражать данную информацию в процентной ставке, что определяется числом классов при классификации физических лиц по кредитоспособности. Используя бинарные методы банк не может дифференцировать процентную ставку на основе результата классификации, потому что клиенты с классом “кредитоспособные” неразличимы между собой. Для иллюстрации указанной особенности приведён пример предоставления клиентами частной информации при использовании банками различных по числу классов методов классификации физических лиц по кредитоспособности. Допустим, что на рынке потребительского кредитования 5 банков, которые применяют методы классификации физических лиц по кредитоспособности с различным числом классов:

1. с одним, кредитный риск не оценивается;

2. с двумя, как в большинстве методов кредитного скоринга;
3. с тремя, подобно работе [80];
4. с пятью, подобно работам [109, 144];
5. с бесконечным числом классов (классификация по непрерывной шкале).

**Действия банков.** Банк устанавливает процентную ставку по кредиту в соответствии с предложенной ценовой стратегией. Функция  $(Q)$  характеризует премию за риск при различном классе кредитоспособности клиента, которая является убывающей функцией. Функция  $P(Q) = P_{min} + C(Q)$  определяет процентную ставку по кредиту при различном классе кредитоспособности клиента, которая так же является убывающей функцией. Заёмщики, принадлежащие к одному классу кредитоспособности, получают одинаковую процентную ставку, соответствующую средней себестоимости кредита для соответствующего класса. Клиентам с наибольшим кредитным риском банк отказывает в выдаче кредита<sup>5</sup>.

**Действия клиентов.** Клиенты самостоятельно решают, какую информацию предоставлять в банк. Информация имеет срок давности. Для того, чтобы информация была актуальной, её необходимо обновлять. Каждый клиент имеет рациональное поведение: он не станет предоставлять частную информацию, если не получит материальной выгоды от банка. 1000 клиентов подают заявки на кредит в каждом банке ( $N = 1000$ ). Отдельно взятый клиент при предоставлении частной информации не замечает изменения средней процентной ставки по группе, в которую он классифицирован. Клиенты упорядочены по уровню кредитоспособности (3.51).

**Анализ информации.** Банк классифицирует клиентов по кредитоспособности на основе предоставляемой информации. Значение допустимого класса кредитоспособности, ниже которого заёмщики считаются некредитоспособными, равно 0,2. Остальные классы установлены на таком уровне, чтобы охватить равное количество клиентов в рамках используемой шкалы. Изначально клиенты не знают, какая информация увеличивает класс кредитоспособности в банке.

---

<sup>5</sup>За исключением банка № 1, где применяется одна градация, потому что клиенты неразличимы. В банке № 5, где применяется непрерывная шкала, отмечается допустимый уровень кредитоспособности, ниже которого клиентам кредит не выдаётся



Однако в процессе взаимодействия с банком заёмщики начинают определять, какая информация увеличивает класс кредитоспособности, что характеризует состояние равновесия.

$$Q_i = i/1000, \quad i = 1, 2, \dots, 1000, \quad Q_i \in (0, 1] \quad (3.51)$$

$$0 \leq \widehat{Q}_i^m \leq Q_i \quad (3.52)$$

$$\widehat{Q}^m = \sum_{i=1}^N \widehat{Q}_i^m \quad (3.53)$$

Банк № 1 не анализирует кредитоспособность, поэтому предоставление частной информации не влияет на процентную ставку. Все заёмщики получают кредит по единой процентной ставке, равной средней стоимости всех запрошенных кредитов (3.54). В данном случае клиенты не имеют материального стимула предоставлять частную информацию. В состоянии равновесия клиенты не предоставляют частную информацию в данный банк.

$$P_1(Q_i) = P_1 = \overline{P(Q_i)}, \quad Q_i \in (0, 1] \quad (3.54)$$

Банк № 2 классифицирует физических лиц по кредитоспособности по шкале с двумя классами: “кредитоспособен”, “некредитоспособен”. Банк выдаёт кредит только “кредитоспособным” заёмщикам по единой процентной ставке, которая определяется, как средняя стоимость кредитов среди класса “кредитоспособных” заёмщиков (3.55).

$$P_2(Q_i) = P_2 = \overline{P(Q_i)}, \quad \text{при } Q_i \in [0, 2, 1] \quad (3.55)$$

“Некредитоспособные” клиенты не заинтересованы предоставлять частную информацию, т.к. это не приведёт к выдаче кредита. “Кредитоспособные” клиенты заинтересованы предоставить информацию на допустимом уровне, чтобы получить кредит. Предоставление частной информации выше уровня, достаточного для получения статуса кредитоспособного заёмщика, не приведёт к улучшению условий кредитования. В состоянии равновесия некредитоспо-

собные клиенты не предоставляют частную информацию, а кредитоспособные клиенты предоставляют информацию на уровне необходимом, чтобы получить оценку “кредитоспособный” и взять кредит (3.56). В состоянии равновесия общая оценка кредитоспособности равна 160,2.

$$\widehat{Q}_i^2 = \begin{cases} 0, & \text{при } Q_i \in (0, 0,2) \\ 0,2, & \text{при } Q_i \in [0,2, 1] \end{cases} \quad (3.56)$$

Банк № 3 классифицирует физических лиц по кредитоспособности по шкале с тремя классами: “высокой”, “низкой” и “плохой” кредитоспособностью. Клиенты с “плохой” кредитоспособностью не получают кредит. Процентная ставка для клиентов устанавливается следующим образом (3.57):

1. клиенты с “высокой” кредитоспособностью получают кредит по низкой процентной ставке, которая равна средней стоимости кредита данного класса.
2. клиенты с “низкой” кредитоспособностью получают кредит по высокой процентной ставке, которая равна средней стоимости кредита данного класса.

$$P_3(\widehat{Q}_i) = \begin{cases} P_3^{lc}, & \text{при } Q_i \in [0,2, 0,6) \\ P_3^{hc}, & \text{при } Q_i \in [0,6, 1] \end{cases} \quad (3.57)$$

Таким образом, процентная ставка для клиентов с низкой кредитоспособностью выше, чем для клиентов с высокой кредитоспособностью (3.58).

$$P_3^{lc} > P_3^{hc} \quad (3.58)$$

Поведение клиентов в банке № 3 описано следующим образом:

1. клиенты, соответствующие категории с плохой оценкой кредитоспособности, не заинтересованы предоставлять частную информацию в банк № 3, аналогично не кредитоспособным клиентам в банке № 2;
2. клиенты с низкой кредитоспособностью заинтересованы предоставлять частную информацию в банк на уровне достаточном, чтобы получить оценку “низкая кредитоспособность”. Предоставлять информацию свыше

данного уровня не принесёт рассматриваемым клиентам материальной выгоды;

3. клиенты с высокой кредитоспособностью заинтересованы предоставлять частную информацию в банк на уровне достаточном, чтобы получить оценку “высокая кредитоспособность”.

В состоянии равновесия клиенты предоставляют частную информацию на уровне, достаточном для получения наибольшей оценки кредитоспособности (3.59). Общая оценка кредитоспособности по предоставленной информации клиентами в банк № 3 равна 320,6.

$$\widehat{Q}_i^3 = \begin{cases} 0, & \text{при } Q_i \in (0, 0,2) \\ 0,2, & \text{при } Q_i \in [0,2, 0,6) \\ 0,6, & \text{при } Q_i \in [0,6, 1] \end{cases} \quad (3.59)$$

Банк № 4 классифицирует физических лиц по кредитоспособности по шкале с пятью классами: “лучшая” (best), “хорошая” (good), “удовлетворительная” (satisfactory), “плохая” (bad) и “худшая” (worst) кредитоспособность. Клиенты с “лучшей”, “хорошей”, “удовлетворительной” и “плохой” кредитоспособностью получают кредит по процентной ставке, равной средней стоимости соответствующей группы (3.60). Процентная ставка для клиентов тем ниже, чем выше оценка кредитоспособности (3.61).

$$P_4(\widehat{Q}_i) = \begin{cases} P_4^{bad}, & \text{при } Q_i \in [0,2, 0,4) \\ P_4^{satisfactory}, & \text{при } Q_i \in [0,4, 0,6) \\ P_4^{good}, & \text{при } Q_i \in [0,6, 0,8) \\ P_4^{best}, & \text{при } Q_i \in [0,8, 1] \end{cases} \quad (3.60)$$

$$P_4^{bad} > P_4^{satisfactory} > P_4^{good} > P_4^{best} \quad (3.61)$$

Клиенты имеют материальный стимул предоставить информацию, чтобы соответствовать как можно более высокому классу кредитоспособности для снижения процентной ставки по кредиту. Однако клиенты не заинтересованы предоставлять частную информацию между пограничными значениями, пото-

му что это не изменит процентную ставку. В состоянии равновесия банк № 4 получает частную информацию от клиентов (3.62) на минимальном уровне класса кредитоспособности. Общая оценка кредитоспособности по информации, полученной от клиентов в состоянии равновесия равна 400,8.

$$\widehat{Q}_i^4 = \begin{cases} 0, & \text{при } \widehat{Q}_i \in (0, 0,2) \\ 0,2, & \text{при } \widehat{Q}_i \in [0,2, 0,4) \\ 0,4, & \text{при } \widehat{Q}_i \in [0,4, 0,6) \\ 0,6, & \text{при } \widehat{Q}_i \in [0,6, 0,8) \\ 0,8, & \text{при } \widehat{Q}_i \in [0,8, 1] \end{cases} \quad (3.62)$$

Банк № 5 применяет непрерывную шкалу для классификации физических лиц по кредитоспособности, что позволяет дифференцировать процентную ставку по непрерывной шкале. Заёмщик с уровнем кредитоспособности выше допустимого заинтересован предоставить всю частную информацию в данный банк, так как любая положительная информация снижает процентную ставку, а негативная информация оставляет ставку неизменной. В состоянии равновесия заёмщик предоставит частную информацию на уровне, который соответствует действительной кредитоспособности. Общая оценка кредитоспособности для банка № 5 равна 480.

В таблице 3.7 показано значение предоставленной клиентами частной информации, оцененная кредитоспособностью, в банках, которые используют различные шкалы для оценки. В рассматриваемом примере заёмщик с уровнем кредитоспособности 0,5 предоставляет частную информацию в банки таким образом, что в банке № 1 оценка кредитоспособности равна 0, в банке № 2 равна 0,2, в банке № 3 равна 0,2, в банке № 4 равна 0,4, в банке № 5 равна 0,5 (Рис. 3.8).

На рисунках 3.8 - 3.9 показана зависимость оценки кредитоспособности, построенной по предоставляемой информации в банк №  $m$  ( $\widehat{Q}_i^m$ ), от действительного уровня кредитоспособности клиента ( $Q_i$ ) с учётом ценовой стратегии, используемой банком. По оси абсцисс  $Q$  отмечен уровень кредитоспособности, которым обладает клиент. По оси ординат  $\widehat{Q}$  отмечена оценка кредитоспособности по информации, которую клиент предоставил в банк. Под наклонной линией находится множество допустимых значений оценки кредитоспособности для

Таблица 3.7: Класс кредитоспособности, вычисленный по предоставленной клиентами информации

$i$	Класс кредитоспособности клиента $i$									$\widehat{Q}^m = \sum_{i=1}^{1000} \widehat{Q}_i^m$
	100	200	300	400	500	600	700	800	900	
Банк № 1 ( $\widehat{Q}_i^1$ )	0	0	0	0	0	0	0	0	0	0
Банк № 2 ( $\widehat{Q}_i^2$ )	0	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	160,2
Банк № 3 ( $\widehat{Q}_i^3$ )	0	0,2	0,2	0,2	0,2	0,6	0,6	0,6	0,6	320,6
Банк № 4 ( $\widehat{Q}_i^4$ )	0	0,2	0,2	0,4	0,4	0,6	0,6	0,8	0,8	400,8
Банк № 5 ( $\widehat{Q}_i^4$ )	0	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	480

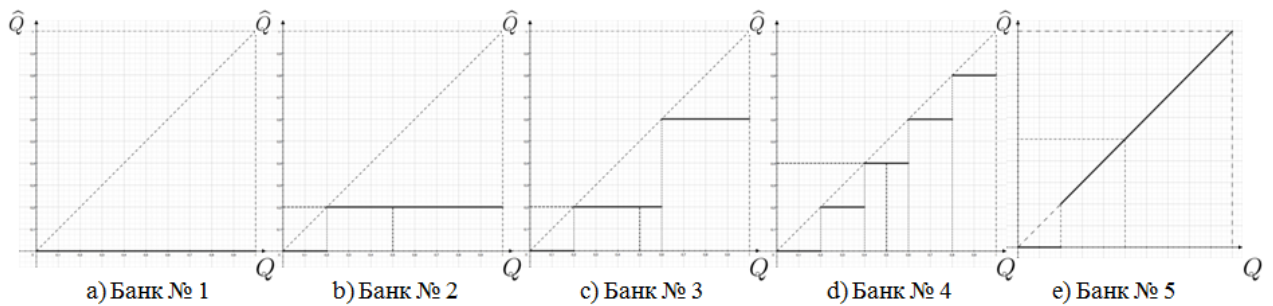


Рисунок 3.8: Класс кредитоспособности клиента в банках при  $Q = 0,5$

клиентов (3.52). Клиент не может получить оценку кредитоспособности выше действительного уровня. Для получения суммарной оценки предоставленной информации клиентами в банк №  $m$  оценки кредитоспособности в данном банке каждого клиента просуммированы, что обозначено, как общая оценка кредитоспособности в банке №  $m$  (3.53).

Для сравнения заёмщик с уровнем кредитоспособности, равным 0,9 предоставляет частную информацию в банки таким образом, что в банке № 1 оценка кредитоспособности равна 0, в банке № 2 равна 0,2, в банке № 3 равна 0,6, в банке № 4 равна 0,8, в банке № 5 равна 0,9 (Рис. 3.9).



Рисунок 3.9: Класс кредитоспособности клиента в банках при  $Q = 0,9$

Таким образом эффективность построенной ценовой стратегии зависит от точности оценки кредитоспособности. Данная точность влияет на возможности банка отразить частную информацию заёмщиков в процентной ставке по кредиту. В продемонстрированном примере влияния количества градаций оценки кредитоспособности на предоставляемую частную информацию клиентами в банк по рассматриваемой ценовой стратегии результаты структурированы следующим образом:

- банк, в котором не дифференцирована процентная ставка, клиенты не заинтересованы предоставлять частную информацию;
- банк, в котором клиенты дифференцированы по дискретной шкале кредитоспособности, клиенты заинтересованы предоставлять частную информацию на уровне, достаточном, чтобы перейти на как можно более высокую ступень кредитоспособности;
- банк, в котором процентная ставка для клиента устанавливается индивидуально (по непрерывной шкале) в соответствии с ожидаемыми потерями по ссудам клиенты заинтересованы предоставить всю положительную частную информацию.

В результате, чем больше классов при классификации физических лиц кредитоспособности при использовании сформулированной ценовой стратегии, тем больше частной информации заёмщики предоставляют в банк (3.63). Применение непрерывной шкалы при классификации физических лиц по кредитоспособности обеспечивает полное отражение информации клиента значимой в процентной ставке по кредиту, что позволяет банку сформировать обратную связь, по которой заёмщик материально заинтересован предоставлять всю частную информацию, чтобы получить более дешёвый кредит. Применение построенной ценовой стратегии с классификацией физических лиц по кредитоспособности по непрерывной шкале ориентировано на стимулирование клиентов банка действовать согласно принципу полного раскрытия на рынках с асимметрией информации. Таким образом, используя предложенную ценовую стратегию, то есть устанавливая изначально наиболее высокую процентную ставку и снижая её только при предоставлении клиентом информации, которая повышает класс

кредитоспособности, банк может снизить расход средств на сбор информации. Стоит отметить, что разработанный метод классификации объектов кредитоспособности по непрерывной шкале позволяет отразить всю значимую измеримую информацию в итоговой оценке, что позволяет реализовать предложенную ценовую стратегию.

$$0 = \widehat{Q}^1 < \widehat{Q}^2 < \widehat{Q}^3 < \widehat{Q}^4 < \widehat{Q}^5 = \sum Q_i \quad (3.63)$$

Ценовая стратегия при выдаче кредита требует дальнейшего изучения, потому что даже в давно применяемой системе ценообразования при кредитовании юридических лиц существуют ситуации, в которых целесообразно отклоняться от установления процентной ставки в соответствии с кредитным риском (risk based pricing). Например, когда организация находится в трудных финансовых условиях, закрытие доступа к кредитным ресурсам (снижение кредитного лимита) может стать причиной для банкротства компании, поэтому в некоторых случаях кредитные агентства завышают класс кредитоспособности фирмы [97]. Аналогично высокие процентные ставки по кредиту физическим лицам могут быть причиной банкротства последних. Так же на рынке капитала участники могут отклоняться от риск-нейтральной стратегии для извлечения спекулятивной выгоды [73].

Стоит отметить, что класс кредитоспособности полученный по предложенной модели может быть использован для оценки риска единицы выданных кредитов в портфеле кредитов, выданных физическим лицам. Для осуществления подобной оценки достаточно класс кредитоспособности клиентов банка в портфеле потребительских кредитов взвесить по величине кредита относительно величины портфеля (3.64).

$$\bar{R} = \sum_{i=1}^N (1 - Q_i) \frac{M_i}{M} \quad (3.64)$$

где  $\bar{R}$  – оценка риска единицы выданных кредитов в портфеле потребительских кредитов,  $Q_i$  – кредитоспособность  $i$ -го клиента (физического лица),  $M_i$  – объем кредита, взятого  $i$ -ым клиентом,  $M$  – объем портфеля кредитов,  $N$  – число клиентов.

## Результаты главы

В текущей главе модифицированный метод рандомизированных сводных показателей был апробирован по базе кредитных историй для классификации физических лиц по кредитоспособности. Разработана система поддержки принятия решений (СППР) по установлению процентной ставки физическим лицам. В отличие от существующих СППР, предназначенных для отсева наименее кредитоспособных клиентов и выдаче кредита остальным по единой процентной ставке, разработанная СППР позволяет банку в автоматизированном режиме дифференцировать процентную ставку в заданном диапазоне в соответствии с индивидуальным классом кредитоспособности. Применение разработанной СППР обеспечит более высокую финансовую стабильность банка в секторе кредитования физических лиц.

Разработана алгоритм по установлению процентной ставки физическим лицам. В отличие от существующих моделей СППР при выдаче кредита, где характеристики физического лица и кредита анализируются вместе, разработанная модель основана на независимом анализе кредитоспособности физического лица от параметров кредитной сделки, что позволяет построить множество подпространств уровня, каждое из которых выражает комбинации параметров кредита с одинаково оплаченным риском для банка при конкретном классе кредитоспособности клиента. Применение разработанной модели позволяет банку визуализировать допустимые варианты кредитования клиентам в соответствии с индивидуальным классом кредитоспособности и управлять атрибутами кредитной линии по динамике кредитоспособности.

Указаны особенности ценообразования кредита при использовании разработанной СППР, которая основана на дифференциации процентной ставки по кредиту физическим лицам в соответствии с классификацией по кредитоспособности по непрерывной шкале. В отличии от наиболее распространённой ценовой стратегии, которая заключается в отсева наименее кредитоспособных клиентов и выдачи кредита остальным по единой процентной ставке, а так же остальных стратегий, основанных на дискретной дифференциации процентной ставки, по предложенной стратегии любое измеримое улучшение характеристик клиента снижает процентную ставку по кредиту, что стимулирует последнего предо-



ставлять частную информацию в банк. Предоставление информации клиентом снижает асимметрию информации, позволяя банку снизить расходы на сбор информации о клиентах, что отразится на снижении процентной ставке по кредиту.

Оценен ожидаемый экономический эффект от применения разработанного метода, который состоит в снижении потерь по ссудам на 4,66% от величины выданных средств. Вычислено, что разработка и внедрение подобной СППР банком окупается при объеме потребительских кредитов в 496,4 млн. рублей. Выявлено, что вероятность верно предсказанных исходов (общая точность классификатора), которая была в основе вычислений указанной выше оценки, не отражает всех преимуществ метода классификации объектов по непрерывной шкале. Проведённый анализ по методу Монте-Карло показал, что переход на разработанную систему поддержки принятия решений снизит вероятность отзыва лицензии банка задействованного, в основном, в секторе кредитования физических лиц на величину от 2,89% до 10,93% в зависимости от инвестирования средств в настройку параметров СППР.

## Заключение

Поставленная цель достигнута, необходимые задачи выполнены. На основании проведённого диссертационного исследования было определено следующее. Кредитоспособность клиента банка может принимать бесконечное множество значений. Для точного отображения кредитоспособности в процентной ставке по кредиту необходимо классифицировать физических лиц по кредитоспособности по непрерывной шкале. Предложенная модель классификации объектов по непрерывной шкале является пределом моделей классификации объектов по дискретной шкале при увеличении числа классов до бесконечности. При использовании предложенной модели при кредитовании физических лиц каждое физическое лицо получает индивидуальный класс кредитоспособности. Любая модель классификации физических лиц по дискретной шкале, в том числе и бинарная, могут быть вычислены посредством преобразования результата полученного из предложенной модели. Указанное свойство обеспечивает полную свободу банка при выборе количества категорий для процентной ставки по кредиту при использовании предложенной модели.

Результаты работы представляют существенную практическую значимость для коммерческих банков. Система поддержки принятия решений на основе метода классификации физических лиц по кредитоспособности по непрерывной шкале расширяет возможности банка в автоматизированном взаимодействии с клиентом по сравнению с методами классификации физических лиц по кредитоспособности по бинарной шкале, например, позволяя дифференцировать процентную ставку по кредиту на требуемое число ценовых групп, визуализировать допустимые варианты кредитования для клиента в соответствии с индивидуальным классом кредитоспособности, а так же управлять атрибутами по открытым кредитным линиям по динамике кредитоспособности заёмщиков.

Более того, с точки зрения кредитного риска банк безразличен к выбору клиента комбинаций условий кредитования, которые соответствуют множеству значений поверхности уровня, построенной системой поддержки принятия решений. Клиент выберет наиболее удобный для себя вариант выплаты кредита. Удобство клиента по кредитным выплатам снижает кредитный риск. Выбирая вариант кредитования, клиент максимизирует свою функцию полезности, поэтому возможность подобного выбора может выступать в качестве фактора для клиента при поиске кредитора. Таким образом, банк применяя разработанную систему может предоставить право выбора условий кредитования клиенту из заданного множества, обеспечивая адаптацию кредитных продуктов относительно каждого клиента, тем самым увеличивая качество банковских услуг.

Стоит отметить социальную значимость полученных результатов. Если банк дифференцирует процентную ставку клиента в соответствии с кредитоспособностью то клиент становится материально заинтересован улучшать свои характеристики для повышения кредитоспособности и получения более дешёвого кредита. Основным критерием социально-экономической эффективности является обеспечение развития личности. Установление индивидуальных условий кредитования будет способствовать оптимальному и справедливому распределению ресурсов, снижая асимметрию информации и стабилизируя сектор потребительского кредитования, что обеспечит физическим лицам более надёжный доступ к ресурсам, предоставляя больше возможностей для развития личности.

Для дальнейшего исследования интерес представляет определение меры, позволяющей сравнивать эффективность классификаторов с различным числом градаций и классификаторов с непрерывной шкалой. Так же необходимо, чтобы разрабатываемые классификаторы учитывали стоимость ошибок различного рода. В качестве направления по развитию метода классификации физических лиц по кредитоспособности по непрерывной шкале является дальнейшее использование методов теории нечётких множеств.

## Список литературы

1. Аналитические показатели : обзор банковского сектора Российской Федерации. – Москва : Центральный Банк РФ, 2015. – № 154. – 75 с.
2. Бабкин А. В. Анализ маркетинговых ситуаций на основе результатов моделирования дискретных переменных / А. В. Бабкин, О. С. Воищева, В. И. Тинякова // Научно-технические ведомости СПбГТУ. – 2006. – №5, т. 2 : Экономические и гуманитарные науки. – С. 256-265.
3. Воронцовский А. В. Современные теории рынка капитала. – Москва: Экономика, 2010. – 719 с.
4. Глазьев С. Ю. Стратегия опережающего развития России в условиях глобального кризиса / С. Ю. Глазьев. – Москва : Экономика, 2010. – 287 с.
5. Глухов В. В. Формы и модели управления финансами в домашнем хозяйстве / В. В. Глухов // Проблемы учёта и финансов. – №1. – 2011.
6. Градов А. П. Экономическая безопасность страны: принципы анализа состояния и противодействия угрозам / А. П. Градов, И. В. Ильин // Экономическая наука современной России. – 2005. – № 3. – С. 88-101.
7. Жук С. Н. Оценка эффективности функционирования сложных систем по иерархической системе показателей / С. Н. Жук // Труды СПИИРАН. – 2013. – Т. 3, № 26. – С. 194-203.
8. Заде Л. А. Основы нового подхода к анализу сложных систем и процессов принятия решений / Л. А. Заде // Математика сегодня : сб. ст. – Москва : Знание, 1974. – С. 5-48.

9. Ильин И. В. Моделирование бизнес-архитектуры проектно-ориентированного предприятия / И. В. Ильин, А. И. Лёвина, А. Р. Антипин // Экономика и управление. – 2013. – № 9. – С. 32-38.
10. Ильин И. В. Вопросы формирования архитектуры инжиниринговых компаний / И. В. Ильин, Ю. Л. Левченко, А. И. Левина // Научно-технические ведомости СПбГПУ. Сер.: Эконом. науки. – 2013. – № 163, т. 2. – С. 48-54.
11. Кобзев В. В. Математическое моделирование производственных систем / В. В. Кобзев, А. Е. Радаев, А. С. Кривченко. – Санкт-Петербург : Изд-во Политехн. ун-та, 2014. – 239 с.
12. Колесов Д. Н. Оценка сложных финансово-экономических объектов с использованием системы поддержки принятия решений АСПИД-3W / Д. Н. Колесов, М. В. Михайлов, Н. В. Хованов. – Санкт-Петербург: ОЦЭМ, 2009. – 69 с.
13. Кузнецов Л. А. Оценка кредитной истории физических лиц на основе нечетких моделей / Л. А. Кузнецов, А. В. Перевозчиков // Управление большими системами. – 2008. – № 21. – С. 84-106.
14. Лотов А. В. Многокритериальные задачи принятия решений / А. В. Лотов, И. И. Поспелова. – Москва : Изд. МГУ им. Ломоносова, 2008. – 197 с.
15. Международная стандартная классификация образования 2011. – Монреаль: Институт статистики ЮНЕСКО, 2013. – 87 с.
16. Некрасова Т. П. Вопрос эффективного управления / Т. П. Некрасова // Вектор региона: образование, наука, производство. – Санкт-Петербург, 2004. – № 2. – С. 68-69.
17. О порядке формирования кредитными организациями резервов на возможные потери по ссудам, по ссудной и приравненной к ней задолженности [Электронный ресурс] : положение Банка России, от 26.03.2004 № 254-П / “Гарант” : информ. прав. портал. – Режим доступа: <http://base.garant.ru/584458/> (4.04.14).

18. Об обязательных нормативах банков : инструкции Банка России № 139-И / под ред. Лунтовского Г. И. // Вестник Банка России. – 2012. – № 74(1392).
19. Отчёт о развитии банковского сектора и банковского надзора в 2012 году / Центральный Банк Российской Федерации. – Москва : Парадиз, 2013. – 120 с.
20. Отчёт о развитии банковского сектора и банковского надзора в 2013 году / Центральный Банк Российской Федерации. – Москва : Парадиз, 2014. – 128 с.
21. Отчёт о развитии банковского сектора и банковского надзора в 2014 году / Центральный Банк Российской Федерации. – Москва : Парадиз, 2015. – 116 с.
22. Пегат А. Нечеткое моделирование и управление / А. Пегат. – Москва: Бинном. Лаборатория знаний. – 2009. – 798 с.
23. Силкина Г. Ю. Инновационные процессы в экономике знаний. Анализ и моделирование / Г. Ю. Силкина, С. Ю. Шевченко. – Санкт-Петербург : Изд. Политехн. ун-та, 2014. – 167 с.
24. Спектр Е. И. О кредитных историях: комментарий к новому закону / Е. И. Спектр // Право и Экономика. – 2005. – № 7. – С. 10-14.
25. Столярова М. В. Оценка кредитоспособности клиентов банка с привлечением аппарата нечеткой математики / М. В. Столярова, Г. Ю. Силкина // XL Неделя науки СПбГПУ. – Санкт-Петербург, 2011. – С. 203-205.
26. Сулоева С. Б. Контроллинг / С. Б. Сулоева, Н. В. Муханова. – Изд. 3-е, испр. и доп. – Санкт-Петербург : Изд-во Политехн. ун-та, 2008 . – 95 с.
27. Федеральный закон от 10.07.2002 № 218-ФЗ “О Центральном Банке Российской Федерации (Банке России)” [Электронный ресурс] // КонсультантПлюс. – Режим доступа: [http : //www.consultant.ru/document/cons\\_doc\\_LAW\\_37570/](http://www.consultant.ru/document/cons_doc_LAW_37570/) (19.03.2016).
28. Федеральный закон от 22.12.2004 № 86-ФЗ “О кредитных историях” [Электронный ресурс] // КонсультантПлюс. – Режим доступа: <http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=168306> (19.03.2016).

29. Юрьев В. Н. Информационные системы в экономике / В. Н. Юрьев, В. Н. Волкова. – Санкт-Петербург : Изд-во Политехн. ун-та, 2006. – 538 с.
30. A case-based reasoning model that uses preference theory functions for credit scoring / S. Vukovic [и др.] // Expert Systems with Applications. – 2012. – Т. 39, № 9 – С. 8389-8395.
31. A comparative assessment of ensemble learning for credit scoring / G. Wang [и др.] // Expert Systems with Applications. – 2011. – Т. 38, № 1. – С. 223-230.
32. A trust-based bio-inspired approach for credit lending decisions / M. S. Mirtalaei [и др.] // Computing. – 2012. – Т. 94, № 7. – С. 541-577.
33. Abdou H. A. Credit scoring, statistical techniques and evaluation criteria: A review of the literature / H. A. Abdou, J. Pointon // Intelligent Systems in Accounting, Finance and Management. – 2011. – Т. 18, № 2-3. – С. 59-88.
34. Akerlof G. A. The market for “lemons”: quality uncertainty and the market mechanism / G. A. Akerlof // The Quarterly Journal of Economics. – 1970. – № 84. – С. 488-500.
35. An application of locally linear model tree algorithm for predictive accuracy of credit scoring / M. Siami [и др.] // Model and Data Engineering. – Berlin Heidelberg: Springer, 2011. – С. 133-142.
36. Analysing user trust in electronic banking using data mining methods / F. Liebana-Cabanillas [и др.] // Expert Systems with Applications. – 2013. – Т. 40, № 14. – С. 5439-5447.
37. Anderson R. The credit scoring toolkit: theory and practice for retail credit risk management and decision automation / R. Anderson. – Oxford: Oxford University Press, 2007. – 731 с.
38. Artificial neural network models for forecasting and decision making / T. Hill [и др.] // International Journal of Forecasting. – 1994. – Т. 10, № 1. – С. 5-15.

39. Assessing the sustainability of the energy use of residential buildings in Belgrade through multi-criteria analysis / Vucicevic B. [и др.] // *Energy and Buildings*. – 2014. – № 69. – С. 51-61.
40. Assessing the global sustainability of different electricity generation systems / J. J. C. Barros // *Energy*. – 2015. – Т. 89. – С. 473-489.
41. Bahrammirzaee A. A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems / A. Bahrammirzaee // *Neural Computing and Applications*. – 2010. – Т. 19, № 8. – С. 1165-1195.
42. Barro R. J. A new data set of educational attainment in the world, 1950-2010 / R. J. Barro, J. W. Lee // *Journal of Development Economics*. – 2013. – № 104. – С. 184-198.
43. Bellotti T. Are rating agencies' assignments opaque? Evidence from international banks / T. Bellotti, R. Matousek, C. Stewart // *Expert Systems with Applications*. – 2011. – Т. 38, № 4. – С. 4206-4214.
44. Bellotti T. Support vector machines for credit scoring and discovery of significant features / T. Bellotti, J. Crook // *Expert Systems with Applications*. – 2009. – Т. 36, № 2. – С. 3302-3308.
45. Benchmarking state-of-the-art classification algorithms for credit scoring / B. Baesens [и др.] // *Journal of the Operational Research Society*. – 2003. – Т. 54, № 6. – С. 627-635.
46. Brown M. Information sharing and credit: Firm-level evidence from transition countries / M. Brown, T. Jappelli // *Journal of Financial Intermediation*. – 2009. – Т. 18, № 2. – С. 151-172.
47. Bubb R. Securitization and moral hazard: Evidence from credit score cutoff rules / R. Bubb, A. Kaufman // *Journal of Monetary Economics*. – 2014. – № 63. – С. 1-18.



48. Buyukkarabacak B. Credit information sharing and banking crises: An empirical investigation / B. Buyukkarabacak, N. Valev // *Journal of Macroeconomics*. – 2012 – T. 34, № 3. – C. 788-800.
49. Capotorti A. Credit scoring on consumer behavior using fuzzy markov model / A. Capotorti, E. Barbanera // *Computational Statistics and Data Analysis*. – T. 56, № 4. – 2012. – C. 981-994.
50. Chang S. Y. An artificial immune classifier for credit scoring analysis / S. Y. Chang, T. Y. Yeh // *Applied Soft Computing*. – 2012. – T. 12, № 2. – C. 611-618.
51. Chen W. Mining the customer credit using hybrid support vector machine technique / W. Chen, C. Ma, L. Ma // *Expert Systems with Applications*. – 2009. – T. 36, № 4. – C. 7611-7616.
52. Chi B. W. A hybrid approach to integrate genetic algorithm into dual scoring model in enhancing the performance of credit scoring model / B. W. Chi, C. C. Hsu // *Expert Systems with Applications*. – 2012. – T. 39, № 3. – C. 2650-2661.
53. Chi B. W. Enhancing credit scoring model performance by a hybrid scoring matrix / B. W. Chi, C. C. Hsu, M. H. Ho // *African Journal of Business Management*. – 2013. – T. 7, № 18. – C. 1791-1805.
54. Comeig I. Information acquisition in SME's relationship lending and the cost of loans / I. Comeig, M. O. Fernandez-Blanco, F. Ramirez // *Journal of Business Research*. – 2015. – T. 68, № 7. – C. 1650-1652.
55. Cox E. P. The optimal number of response alternatives for a scale: A review / E. P. Cox // *Journal of Marketing Research*. – 1980. – C. 407-422.
56. Credit rating analysis with support vector machines and neural networks: A market comparative study / Z. Huang [и др.] // *Decision Support Systems*. – 2004. – T. 37, № 4 – C. 543-558.
57. Credit risk evaluation using a weighted least squares SVM classifier with design of experiment for parameter selection / L. Yu [и др.] // *Expert Systems with Applications*. – 2011. – T. 38, № 12. – C. 15392-15399.

58. Crook J. N. Recent developments in consumer credit risk assessment / J. N. Crook, D. B. Edelman, L. C. Thomas // *European Journal of Operational Research*. – 2007. – Т. 183, № 3. – С. 1447-1465.
59. Danenas P. Selection of support vector machines based classifiers for credit risk domain / P. Danenas, G. Garsva // *Expert Systems with Applications*. – 2015. – Т. 42, № 6. – С. 3194-3204.
60. Decancq K. Weights in multidimensional indices of wellbeing: An overview / K. Decancq, M. A. Lugo // *Econometric Reviews*. – 2013. – Т. 32, № 1. – С. 7-34.
61. Defining a novel k-nearest neighbours approach to assess the applicability domain of a QSAR model for reliable predictions / F. Sahigara [и др.] // *Journal of Cheminformatics*. – 2013. – № 5. – С. 27.
62. Desai V. S. A comparison of neural networks and linear scoring models in the credit union environment / V. S. Desai, J. N. Crook, G. A. Overstreet // *European Journal of Operational Research*. – 1996. – Т. 95, № 1. – С. 24-37.
63. Dimitrov V. Impact of the Dodd-Frank act on credit ratings / V. Dimitrov, D. Palia, L. Tang // *Journal of Financial Economics*. – 2015. – Т. 115, № 3. – С. 505-520.
64. Dong G. N. Mortgage securitization, housing market and real output: a time-series causality test using structural VAR / G.N. Dong. – Columbia: Columbia University, 2012. – 37 с.
65. Durand D. Risk elements in consumer installment financing / D. Durand. – NY: National Bureau of Economic Research, 1941. – 101 с.
66. Einav L. The impact of credit scoring on consumer lending / L. Einav, M. Jenkins, J. Levin // *The RAND Journal of Economics*. – 2013. – Т. 44, № 2. – С. 249-274.
67. Elyasiani E. Relationship lending: a survey of the literature / E. Elyasiani, L. G. Goldberg // *Journal of Economics and Business*. – 2004. – Т. 56, № 4. – С. 315-330.

68. Fama E. F. Efficient capital markets: A review of theory and empirical work\* / E. F. Fama // The Journal of Finance. – 1970. – T. 25, № 2. – C. 383-417.
69. Fama E. F. Efficient capital markets: 2 / E. F. Fama // The Journal of Finance. – 1991. – T. 46, № 5. – C. 1575-1617.
70. Feraud R. A methodology to explain neural network classification / R. Feraud, F. Clerot // Neural Networks. – 2002. – T. 15, № 2. – C. 237-246.
71. Gaiotti E. Credit availability and investment: Lessons from the “great recession” / E. Gaiotti // European Economic Review. – 2013. – № 59. – C. 212-227.
72. Garcia V. Non-parametric statistical analysis of machine learning methods for credit scoring / V. Garcia, A. I. Marques, J. S. Sanchez // Management Intelligent Systems. – Berlin Heidelberg: Springer, 2012. – C. 263-272.
73. Gloukhov V. V. Market Risk Neutral Strategies: Modeling and Algorithmization / V. V. Gloukhov, I. V. Ilin, V. I. Koposov, A. I. Levina // Asian Social Science. – 2014. – T. 10, № 24. – C. 209-216.
74. Goel A. M. Information reliability and welfare: A theory of coarse credit ratings / A. M. Goel, A. V. Thakor // Journal of Financial Economics. – 2015. – T. 115, № 3. – C. 541-557.
75. Grossman S. J. On the impossibility of informationally efficient markets / S. J. Grossman, J. E. Stiglitz // The American Economic Review. – 1980. – C. 393-408.
76. Hand D. J. Choosing k for two-class nearest neighbour classifiers with unbalanced classes / D. J. Hand, V. Vinciotti // Pattern Recognition Letters. – 2003. – T. 24, № 9. – C. 1555-1562.
77. Hand D. J. Classifier technology and the illusion of progress / D. J. Hand // Statistical Science. – 2006. – T. 21, № 1. – C. 1-14.
78. Hand D. J. Statistical classification methods in consumer credit scoring: a review / D. J. Hand, W. E. Henley // Journal of the Royal Statistical Society. Series A (Statistics in Society). – 1997. – C. 523-541.

79. Harris T. Credit scoring using the clustered support vector machine / T. Harris // Expert Systems with Applications. – 2015. – Т. 42, № 2. – С. 741-750.
80. Hatzilygeroudis I. Fuzzy and Neuro-Symbolic Approaches in Personal Credit Scoring: Assessment of Bank Loan Applicants / I. Hatzilygeroudis, J. Prentzas // Innovations in Intelligent Machines-4. – Springer International Publishing, 2014. – С. 319-339.
81. Henley W. E. A k-nearest-neighbour classifier for assessing consumer credit risk / W. E. Henley, D. J. Hand // The Statistician. – 1996. – С. 77-95.
82. Heuson A. Credit scoring and mortgage securitization: Implications for mortgage rates and credit availability / A. Heuson, W. Passmore, R. Sparks // The Journal of Real Estate Finance and Economics. – 2001. – Т. 23, № 3. – С. 337-363.
83. Hovanov N. Multicriteria estimation of probabilities on basis of expert non-numeric, non-exact and non-complete knowledge / N. Hovanov, M. Yudaeva, K. Hovanov // European Journal of Operational Research. – 2009. – Т. 195, № 3. – С. 857-863.
84. Huang C. L. Credit scoring with a data mining approach based on support vector machines / C. L. Huang, M. C. Chen, C. J. Wang // Expert Systems with Applications. – 2007. – Т. 33, № 4. – С. 847-856.
85. Huang S. C. Using Gaussian process based kernel classifiers for credit rating forecasting / S. C. Huang // Expert Systems with Applications. – 2011. – Т. 38, № 7. – С. 8607-8611.
86. Investigation and improvement of multi-layer perception neural networks for credit scoring / Z. Zhao [и др.] // Expert Systems with Applications. – 2015. – Т. 42, № 7. – С. 3508-3516.
87. Jappelli T. Information sharing, lending and defaults: Cross-country evidence / T. Jappelli, M. Pagano // Journal of Banking and Finance. – 2002. – Т. 26, № 10. – С. 2017-2045.

88. Johnson K. W. Recent developments in the credit card market and the financial obligations ratio / K. W. Johnson // *Federal Reserve Bulletin*. – 2005. – № 91. – C. 473.
89. Ju Y. Behavioral technology credit scoring model with time-dependent rates for stress test / Y. Ju, S. Y. Jeon, S. Y. Sohn // *European Journal of Operational Research*. – 2015. – T. 242, № 3. – C. 910-919.
90. Karlan D. Observing unobservables: Identifying information asymmetries with a consumer credit field experiment / D. Karlan, J. Zinman // *Econometrica*. – 2009. – T. 77, № 6. – C. 1993-2008.
91. Kau J. B. Racial discrimination and mortgage lending / J. B. Kau, D. C. Keenan, H. J. Munneke // *The Journal of Real Estate Finance and Economics*. – 2012. – T. 45, № 2. – C. 289-304.
92. Khashman A. Credit risk evaluation using neural networks: Emotional versus conventional models / A. Khashman // *Applied Soft Computing*. – 2011. – T. 11, № 8. – C. 5477-5484.
93. Kim G. Understanding dynamics between initial trust and usage intentions of mobile banking / G. Kim, B. S. Shin, H. G. Lee // *Information Systems Journal*. – 2009. – T. 19, № 3. – C. 283-311.
94. Kim Y. S. Managing loan customers using misclassification patterns of credit scoring model / Y. S. Kim, S. Y. Sohn // *Expert Systems with Applications*. – 2004. – T. 26, № 4. – C. 567-573.
95. Korkeamaki T. Credit ratings and information asymmetry on the Chinese syndicated loan market / T. Korkeamaki, S. Poyry, M. Suo // *China Economic Review*. – 2014. – № 31. – C. 1-16.
96. Kozeny V. Genetic algorithms for credit scoring / V. Kozeny // *Expert Systems with Applications: An International Journal*. – 2015. – T. 42, № 6. – C. 2998-3004.
97. Kraft P. Do rating agencies cater? Evidence from rating-based contracts / P. Kraft // *Journal of Accounting and Economics*. – 2015. – T. 59, № 2. – C. 264-283.

98. Laffont J. J. The theory of incentives: the principal-agent model / J. J. Laffont, D. Martimort. – Princeton: Princeton University Press, 2001. – 430 c.
99. Laha A. Developing credit scoring models with SOM and fuzzy rule based k-NN classifiers / A. Laha // IEEE International Conference on Fuzzy Systems, 2006. – IEEE, 2006. – C. 692-698.
100. Lee C. F. Encyclopedia of finance / C. F. Lee, A. C. Lee. – Springer Science and Business Media, 2006. – 855 c.
101. Li S. Research on Impact of Moral Hazard on Individual Credit Risk / S. Li, Y. Yang, Z. Zongfang // Procedia Computer Science. – 2014. – № 31. – C. 577-586.
102. Li X. Why do we trust new technology? A study of initial trust formation with organizational information systems / X. Li, T. J. Hess, J. S. Valacich // The Journal of Strategic Information Systems. – 2008. – T. 17, № 1. – C. 39-71.
103. Lin W. Y. Machine learning in financial crisis prediction: a survey / W. Y. Lin, Y. H. Hu, C. F. Tsai // IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews. – IEEE, 2012. – T. 42, № 4. – C. 421-436.
104. Liu K. Dynamic Credit Scoring on Consumer Behavior Using Fuzzy Markov Model / K. Liu, K. K. Lai, S. M. Guu // Fourth International Multi-Conference on Computing in the Global Information Technology, 2009. – IEEE, 2009. – C. 235-239.
105. Luo X. Examining multi-dimensional trust and multi-faceted risk in initial acceptance of emerging technologies: An empirical study of mobile banking services / X. Luo, H. Li, J. Zhang, J. P. Shim // Decision Support Systems. – 2010. – T. 49, № 2. – C. 222-234.
106. Magri S. The rise of risk-based pricing of mortgage interest rates in Italy / S. Magri, R. Pico // Journal of Banking and Finance. – 2011. – T. 35, № 5. – C. 1277-1290.
107. Malaquias R. F. An empirical study on trust in mobile banking: A developing country perspective / R. F. Malaquias, Y. Hwang // Computers in Human Behavior. – 2016. – № 54. – C. 453-461.

108. Malhotra R. Differentiating between good credits and bad credits using neuro-fuzzy systems / R. Malhotra, D. K. Malhotra // European Journal of Operational Research. – 2002. – T. 136, № 1. – C. 190-211.
109. Marikkannu P. Classification of customer credit data for intelligent credit scoring system using fuzzy set and MC2—Domain driven approach / P. Marikkannu, K. Shanmugapriya // 3rd International Conference on Electronics Computer Technology, 2011. – IEEE, 2011. – № 3. – C. 410-414.
110. Marques A. I. A literature review on the application of evolutionary computing to credit scoring / A. I. Marques, V. Garcia, J. S. Sanchez // Journal of the Operational Research Society. – 2012. – T. 64, № 9. – C. 1384-1399.
111. Martens D. Comprehensible credit scoring models using rule extraction from support vector machines / D. Martens, B. Baesens, T. Van Gestel, J. Vanthienen // European Journal of Operational Research. – 2007. – T. 183, № 3. – C. 1466-1476.
112. Mileris R. Estimates of Loan Applicants' Default Probability Applying Discriminant Analysis and Simple Bayesian Classifier / R. Mileris // Economics and Management. – 2010. – C. 1078-1084.
113. Naeem S. Credit Risk Scorecards Developing and Implementing Intelligent Credit Scoring / S. Naeem. – 2006. – 196 c.
114. Oh J. H. Credit lender-borrower relationship in the credit card market—Implications for credit risk management strategy and relationship marketing / J. H. Oh, W. J. Johnston // International Business Review. – 2014. – T. 23, № 6. – C. 1086-1095.
115. Oreski S. Genetic algorithm-based heuristic for feature selection in credit risk assessment / S. Oreski, G. Oreski // Expert Systems with Applications. – 2014. – T. 41, № 4. – C. 2052-2064.
116. Oreski S. Hybrid system with genetic algorithm and artificial neural networks and its application to retail credit risk assessment / S. Oreski, D. Oreski, G.

- Oreski // Expert Systems with Applications. – 2012. – Т. 39, № 16. – С. 12605-12617.
117. Padilla A. J. Sharing default information as a borrower discipline device / A. J. Padilla, M. Pagano // European Economic Review. – 2000. – Т. 44, № 10. – С. 1951-1980.
118. Reichert A. K. An examination of the conceptual issues involved in developing credit-scoring models / A. K. Reichert, C. C. Cho, G. M. Wagner // Journal of Business and Economic Statistics. – 1983. – Т. 1, № 2. – С. 101-114.
119. Plantin G. Good securitization, bad securitization / G. Plantin. – Tokyo: Institute for Monetary and Economic Studies, Bank of Japan, 2011. – № 11-E-04. – 21 с.
120. Plantin G. Shadow banking and bank capital regulation / G. Plantin // Review of Financial Studies. – 2014. – 49 с.
121. Rohner R. J. Equal Credit Opportunity Act / R. J. Rohner // The Business Lawyer. – 1979. – С. 1423-1433.
122. Rotchild M. Equilibrium in competitive insurance markets / M. Rotchild, J. Stiglitz // Quarterly Journal of Economics. – 1976. – Т. 90, № 4. – С. 629-649.
123. Principles for the Management of Credit Risk / R. Cole [и др.]. – Basel: Basel Committee on Banking Supervision, 2000. – 27 с.
124. Setiono R. Rule extraction from neural networks and support vector machines for credit scoring / R. Setiono, B. Baesens, D. Martens // Data Mining: Foundations and Intelligent Paradigms. – Berlin Heidelberg: Springer, 2012. – С. 299-320.
125. Shmelev S. E. Dynamic sustainability assessment: The case of Russia in the period of transition (1985–2008) / S. E. Shmelev // Ecological Economics. – 2011. – Т. 70, № 11. – С. 2039-2049.



126. Sohn S. Y. Behavioral credit scoring model for technology-based firms that considers uncertain financial ratios obtained from relationship banking / S. Y. Sohn, Y. S. Kim // *Small Business Economics*. – 2013. – Т. 41, № 4. – С. 931-943.
127. Spence M. Job market signaling / M. Spence // *The Quarterly Journal of Economics*. – 1973. – № 87. – С. 355-374.
128. Stiglitz J. E. Credit rationing in markets with imperfect information / J. E. Stiglitz, A. Weiss // *The American Economic Review*. – 1981. – Т. 71, № 3. С. 393-410.
129. Stiglitz J. E. Financial markets and development / J. E. Stiglitz // *Oxford Review of Economic Policy*. – 1989. – Т. 5, № 4. – С. 55-68.
130. Sustainability assessment of residential buildings by non-linear normalization procedure / B. Vucicevic [и др.] // *Energy and Buildings*. – 2013. – Т. 58 – С. 348-354.
131. Tang T. T. Information asymmetry and firms' credit market access: Evidence from Moody's credit rating format refinement / T. T. Tang // *Journal of Financial Economics*. – 2009. – Т. 93, № 2. – С. 325-351.
132. The application of brute force logistic regression to corporate credit scoring models: Evidence from Serbian financial statements / N. Nikolic [и др.] // *Expert Systems with Applications*. – 2013. – Т. 40, № 15. – С. 5932-5944.
133. Thomas L. C. A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers / L. C. Thomas // *International Journal of Forecasting*. – 2000. – Т. 16, № 2. – С. 149-172.
134. Thomas L. C. Consumer finance: Challenges for operational research / L. C. Thomas // *Journal of the Operational Research Society*. – 2010. – Т. 61, № 1. – С. 41-52.
135. Thomas L. C. *Consumer Credit Models: Pricing, Profit and Portfolios*. / L. C. Thomas – Oxford: OUP, 2009. – 385 с.

136. Tomczak J. M. Classification Restricted Boltzmann Machine for comprehensible credit scoring model / J. M. Tomczak, M. Zięba // Expert Systems with Applications. – 2015. – Т. 42, № 4. – С. 1789-1796.
137. Torregrosa D. Interest Rate Differentials Between, Jumbo and Conforming Mortgages, 1995-2000 / D. Torregrosa. – Washington DC: Congressional Budget Office (US Congress), 2001. – 55 с.
138. Tsai C. F. Using neural network ensembles for bankruptcy prediction and credit scoring / C. F. Tsai, J. W. Wu // Expert Systems with Applications. – 2008. – Т. 34, № 4. – С. 2639-2649.
139. Two credit scoring models based on dual strategy ensemble trees / G. Wang [и др.] // Knowledge-Based Systems. – 2012. – № 26. – С. 61-68.
140. Wang Z. Active learning with adaptive regularization / Z. Wang, S. Yan, C. Zhang // Pattern Recognition. – 2011. – Т. 44, № 10. – С. 2375-2383.
141. West D. Neural network credit scoring models / D. West // Computers and Operations Research. – 2000. – Т. 27, № 11. – С. 1131-1152.
142. When does information asymmetry affect the cost of capital? / C. S. Armstrong [и др.] // Journal of Accounting Research. – 2011. – Т. 49, № 1. – С. 1-40.
143. Xiao W. A comparative study of data mining methods in consumer loans credit scoring management / W. Xiao, Q. Zhao, Q. Fei // Journal of Systems Science and Systems Engineering. – 2006. – Т. 15, № 4. – С. 419-435.
144. Xinhui C. On Consumer Credit Scoring Based on Multi-criteria Fuzzy Logic / C. Xinhui, Q. Zhong // International Conference on Business Intelligence and Financial Engineering, 2009. – IEEE, 2009. – С. 765-768.
145. Yu L. An intelligent-agent-based fuzzy group decision making model for financial multicriteria decision support: The case of credit scoring / L. Yu, S. Wang, K. K. Lai // European Journal of Operational Research. – 2009. – Т. 195, № 3. – С. 942-959.

146. Yu L. Credit risk assessment with a multistage neural network ensemble learning approach / L. Yu, S. Wang, K. K. Lai // Expert Systems with Applications. – 2008. – T. 34, № 2. – C. 1434-1444.
147. Zadeh L. A. Fuzzy logic = computing with words / L. A. Zadeh // Fuzzy Systems, IEEE Transactions on. – 1996. – T. 4, № 2. – C. 103-111.
148. Zadeh L. A. Fuzzy sets / L. A. Zadeh // Information and Control. – 1965. – T. 8, № 3. – C. 338-353.
149. Zhou X. Y. A new credit scoring method based on rough sets and decision tree / X. Y. Zhou, D. F. Zhang, Y. Jiang // Advances in Knowledge Discovery and Data Mining. – Berlin Heidelberg: Springer, 2008. – C. 1081-1089.
150. Zurada J. Performance assessment of data mining methods for loan granting decisions: a preliminary study / J. Zurada, N. Kunene // Artificial Intelligence and Soft Computing. – Berlin Heidelberg: Springer, 2010. – C. 495-502.

## Приложение А.

### Вычисления класса кредитоспособности

В данном приложении приведены исходные данные, промежуточные вычисления, а так же искомая классификация физических лиц по кредитоспособности.

#### А.1 Исходные данные

Первые 100 записей, используемой базы кредитных историй.

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
1	A11	6	A34	A43	1169	A65	A75	4	A93	A101
2	A12	48	A32	A43	5951	A61	A73	2	A92	A101
3	A14	12	A34	A46	2096	A61	A74	2	A93	A101
4	A11	42	A32	A42	7882	A61	A74	2	A93	A103
5	A11	24	A33	A40	4870	A61	A73	3	A93	A101
6	A14	36	A32	A46	9055	A65	A73	2	A93	A101
7	A14	24	A32	A42	2835	A63	A75	3	A93	A101
8	A12	36	A32	A41	6948	A61	A73	2	A93	A101
9	A14	12	A32	A43	3059	A64	A74	2	A91	A101
10	A12	30	A34	A40	5234	A61	A71	4	A94	A101
11	A12	12	A32	A40	1295	A61	A72	3	A92	A101
12	A11	48	A32	A49	4308	A61	A72	3	A92	A101
13	A12	12	A32	A43	1567	A61	A73	1	A92	A101
14	A11	24	A34	A40	1199	A61	A75	4	A93	A101
15	A11	15	A32	A40	1403	A61	A73	2	A92	A101
16	A11	24	A32	A43	1282	A62	A73	4	A92	A101
17	A14	24	A34	A43	2424	A65	A75	4	A93	A101
18	A11	30	A30	A49	8072	A65	A72	2	A93	A101
19	A12	24	A32	A41	12579	A61	A75	4	A92	A101
20	A14	24	A32	A43	3430	A63	A75	3	A93	A101
21	A14	9	A34	A40	2134	A61	A73	4	A93	A101
22	A11	6	A32	A43	2647	A63	A73	2	A93	A101

*продолжение следует*

<i>(продолжение)</i>										
	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
23	A11	10	A34	A40	2241	A61	A72	1	A93	A101
24	A12	12	A34	A41	1804	A62	A72	3	A93	A101
25	A14	10	A34	A42	2069	A65	A73	2	A94	A101
26	A11	6	A32	A42	1374	A61	A73	1	A93	A101
27	A14	6	A30	A43	426	A61	A75	4	A94	A101
28	A13	12	A31	A43	409	A64	A73	3	A92	A101
29	A12	7	A32	A43	2415	A61	A73	3	A93	A103
30	A11	60	A33	A49	6836	A61	A75	3	A93	A101
31	A12	18	A32	A49	1913	A64	A72	3	A94	A101
32	A11	24	A32	A42	4020	A61	A73	2	A93	A101
33	A12	18	A32	A40	5866	A62	A73	2	A93	A101
34	A14	12	A34	A49	1264	A65	A75	4	A93	A101
35	A13	12	A32	A42	1474	A61	A72	4	A92	A101
36	A12	45	A34	A43	4746	A61	A72	4	A93	A101
37	A14	48	A34	A46	6110	A61	A73	1	A93	A101
38	A13	18	A32	A43	2100	A61	A73	4	A93	A102
39	A13	10	A32	A44	1225	A61	A73	2	A93	A101
40	A12	9	A32	A43	458	A61	A73	4	A93	A101
41	A14	30	A32	A43	2333	A63	A75	4	A93	A101
42	A12	12	A32	A43	1158	A63	A73	3	A91	A101
43	A12	18	A33	A45	6204	A61	A73	2	A93	A101
44	A11	30	A34	A41	6187	A62	A74	1	A94	A101
45	A11	48	A34	A41	6143	A61	A75	4	A92	A101
46	A14	11	A34	A40	1393	A61	A72	4	A92	A101
47	A14	36	A32	A43	2299	A63	A75	4	A93	A101
48	A11	6	A32	A41	1352	A63	A71	1	A92	A101
49	A14	11	A34	A40	7228	A61	A73	1	A93	A101
50	A14	12	A32	A43	2073	A62	A73	4	A92	A102
51	A12	24	A33	A42	2333	A65	A72	4	A93	A101
52	A12	27	A33	A41	5965	A61	A75	1	A93	A101
53	A14	12	A32	A43	1262	A61	A73	3	A93	A101
54	A14	18	A32	A41	3378	A65	A73	2	A93	A101
55	A12	36	A33	A40	2225	A61	A75	4	A93	A101
56	A14	6	A31	A40	783	A65	A73	1	A93	A103
57	A12	12	A32	A43	6468	A65	A71	2	A93	A101
58	A14	36	A34	A43	9566	A61	A73	2	A92	A101
59	A13	18	A32	A40	1961	A61	A75	3	A92	A101
60	A11	36	A34	A42	6229	A61	A72	4	A92	A102
61	A12	9	A32	A49	1391	A61	A73	2	A94	A101
62	A12	15	A34	A43	1537	A65	A75	4	A93	A103
63	A12	36	A30	A49	1953	A61	A75	4	A93	A101
64	A12	48	A30	A49	14421	A61	A73	2	A93	A101
65	A14	24	A32	A43	3181	A61	A72	4	A92	A101
66	A14	27	A32	A45	5190	A65	A75	4	A93	A101

*продолжение следует*

<i>(продолжение)</i>										
	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
67	A14	12	A32	A43	2171	A61	A72	2	A92	A101
68	A12	12	A32	A40	1007	A64	A73	4	A94	A101
69	A14	36	A32	A46	1819	A61	A73	4	A93	A101
70	A14	36	A32	A43	2394	A65	A73	4	A92	A101
71	A14	36	A32	A41	8133	A61	A73	1	A92	A101
72	A14	7	A34	A43	730	A65	A75	4	A93	A101
73	A11	8	A34	A410	1164	A61	A75	3	A93	A101
74	A12	42	A34	A49	5954	A61	A74	2	A92	A101
75	A11	36	A32	A46	1977	A65	A75	4	A93	A101
76	A11	12	A34	A41	1526	A61	A75	4	A93	A101
77	A11	42	A32	A43	3965	A61	A72	4	A93	A101
78	A12	11	A33	A43	4771	A61	A74	2	A93	A101
79	A14	54	A30	A41	9436	A65	A73	2	A93	A101
80	A12	30	A32	A42	3832	A61	A72	2	A94	A101
81	A14	24	A32	A43	5943	A65	A72	1	A92	A101
82	A14	15	A32	A43	1213	A63	A75	4	A93	A101
83	A14	18	A32	A49	1568	A62	A73	3	A92	A101
84	A11	24	A32	A410	1755	A61	A75	4	A92	A103
85	A11	10	A32	A43	2315	A61	A75	3	A93	A101
86	A14	12	A34	A49	1412	A61	A73	4	A92	A103
87	A12	18	A34	A42	1295	A61	A72	4	A92	A101
88	A12	36	A32	A46	12612	A62	A73	1	A93	A101
89	A11	18	A32	A40	2249	A62	A74	4	A93	A101
90	A11	12	A30	A45	1108	A61	A74	4	A93	A101
91	A14	12	A34	A43	618	A61	A75	4	A93	A101
92	A11	12	A34	A41	1409	A61	A75	4	A93	A101
93	A14	12	A34	A43	797	A65	A75	4	A92	A101
94	A13	24	A34	A42	3617	A65	A75	4	A93	A102
95	A12	12	A32	A40	1318	A64	A75	4	A93	A101
96	A12	54	A30	A49	15945	A61	A72	3	A93	A101
97	A14	12	A34	A46	2012	A65	A74	4	A92	A101
98	A12	18	A32	A49	2622	A62	A73	4	A93	A101
99	A12	36	A34	A43	2337	A61	A75	4	A93	A101
100	A12	20	A33	A41	7057	A65	A74	3	A93	A101

	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20
1	4	A121	67	A143	A152	2	A173	1	A192	A201
2	2	A121	22	A143	A152	1	A173	1	A191	A201
3	3	A121	49	A143	A152	1	A172	2	A191	A201
4	4	A122	45	A143	A153	1	A173	2	A191	A201
5	4	A124	53	A143	A153	2	A173	2	A191	A201

*продолжение следует*

<i>(продолжение)</i>										
	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20
6	4	A124	35	A143	A153	1	A172	2	A192	A201
7	4	A122	53	A143	A152	1	A173	1	A191	A201
8	2	A123	35	A143	A151	1	A174	1	A192	A201
9	4	A121	61	A143	A152	1	A172	1	A191	A201
10	2	A123	28	A143	A152	2	A174	1	A191	A201
11	1	A123	25	A143	A151	1	A173	1	A191	A201
12	4	A122	24	A143	A151	1	A173	1	A191	A201
13	1	A123	22	A143	A152	1	A173	1	A192	A201
14	4	A123	60	A143	A152	2	A172	1	A191	A201
15	4	A123	28	A143	A151	1	A173	1	A191	A201
16	2	A123	32	A143	A152	1	A172	1	A191	A201
17	4	A122	53	A143	A152	2	A173	1	A191	A201
18	3	A123	25	A141	A152	3	A173	1	A191	A201
19	2	A124	44	A143	A153	1	A174	1	A192	A201
20	2	A123	31	A143	A152	1	A173	2	A192	A201
21	4	A123	48	A143	A152	3	A173	1	A192	A201
22	3	A121	44	A143	A151	1	A173	2	A191	A201
23	3	A121	48	A143	A151	2	A172	2	A191	A202
24	4	A122	44	A143	A152	1	A173	1	A191	A201
25	1	A123	26	A143	A152	2	A173	1	A191	A202
26	2	A121	36	A141	A152	1	A172	1	A192	A201
27	4	A123	39	A143	A152	1	A172	1	A191	A201
28	3	A121	42	A143	A151	2	A173	1	A191	A201
29	2	A121	34	A143	A152	1	A173	1	A191	A201
30	4	A124	63	A143	A152	2	A173	1	A192	A201
31	3	A121	36	A141	A152	1	A173	1	A192	A201
32	2	A123	27	A142	A152	1	A173	1	A191	A201
33	2	A123	30	A143	A152	2	A173	1	A192	A201
34	4	A124	57	A143	A151	1	A172	1	A191	A201
35	1	A122	33	A141	A152	1	A174	1	A192	A201
36	2	A122	25	A143	A152	2	A172	1	A191	A201
37	3	A124	31	A141	A153	1	A173	1	A192	A201
38	2	A121	37	A142	A152	1	A173	1	A191	A201
39	2	A123	37	A143	A152	1	A173	1	A192	A201
40	3	A121	24	A143	A152	1	A173	1	A191	A201
41	2	A123	30	A141	A152	1	A174	1	A191	A201
42	1	A123	26	A143	A152	1	A173	1	A192	A201
43	4	A121	44	A143	A152	1	A172	2	A192	A201
44	4	A123	24	A143	A151	2	A173	1	A191	A201
45	4	A124	58	A142	A153	2	A172	1	A191	A201
46	4	A123	35	A143	A152	2	A174	1	A191	A201
47	4	A123	39	A143	A152	1	A173	1	A191	A201
48	2	A122	23	A143	A151	1	A171	1	A192	A201
49	4	A122	39	A143	A152	2	A172	1	A191	A201

*продолжение следует*

<i>(продолжение)</i>										
	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20
50	2	A121	28	A143	A152	1	A173	1	A191	A201
51	2	A122	29	A141	A152	1	A172	1	A191	A201
52	2	A123	30	A143	A152	2	A174	1	A192	A201
53	2	A123	25	A143	A152	1	A173	1	A191	A201
54	1	A122	31	A143	A152	1	A173	1	A192	A201
55	4	A124	57	A141	A153	2	A173	1	A192	A201
56	2	A121	26	A142	A152	1	A172	2	A191	A201
57	1	A124	52	A143	A152	1	A174	1	A192	A201
58	2	A123	31	A142	A152	2	A173	1	A191	A201
59	2	A123	23	A143	A152	1	A174	1	A191	A201
60	4	A124	23	A143	A151	2	A172	1	A192	A201
61	1	A121	27	A141	A152	1	A173	1	A192	A201
62	4	A121	50	A143	A152	2	A173	1	A192	A201
63	4	A124	61	A143	A153	1	A174	1	A192	A201
64	2	A123	25	A143	A152	1	A173	1	A192	A201
65	4	A122	26	A143	A152	1	A173	1	A192	A201
66	4	A122	48	A143	A152	4	A173	2	A192	A201
67	2	A123	29	A141	A152	1	A173	1	A191	A201
68	1	A121	22	A143	A152	1	A173	1	A191	A201
69	4	A124	37	A142	A153	1	A173	1	A192	A201
70	4	A123	25	A143	A152	1	A173	1	A191	A201
71	2	A122	30	A141	A152	1	A173	1	A191	A201
72	2	A122	46	A143	A151	2	A172	1	A192	A201
73	4	A124	51	A141	A153	2	A174	2	A192	A201
74	1	A121	41	A141	A152	2	A172	1	A191	A201
75	4	A124	40	A143	A152	1	A174	1	A192	A201
76	4	A124	66	A143	A153	2	A174	1	A191	A201
77	3	A123	34	A143	A152	1	A173	1	A191	A201
78	4	A122	51	A143	A152	1	A173	1	A191	A201
79	2	A122	39	A143	A152	1	A172	2	A191	A201
80	1	A122	22	A143	A152	1	A173	1	A191	A201
81	1	A123	44	A143	A152	2	A173	1	A192	A201
82	3	A122	47	A142	A152	1	A173	1	A192	A201
83	4	A122	24	A143	A151	1	A172	1	A191	A201
84	4	A121	58	A143	A152	1	A172	1	A192	A201
85	4	A121	52	A143	A152	1	A172	1	A191	A201
86	2	A121	29	A143	A152	2	A174	1	A192	A201
87	1	A122	27	A143	A152	2	A173	1	A191	A201
88	4	A124	47	A143	A153	1	A173	2	A192	A201
89	3	A123	30	A143	A152	1	A174	2	A192	A201
90	3	A121	28	A143	A152	2	A173	1	A191	A201
91	4	A121	56	A143	A152	1	A173	1	A191	A201
92	3	A121	54	A143	A152	1	A173	1	A191	A201
93	3	A122	33	A141	A152	1	A172	2	A191	A201

*продолжение следует*



<i>(продолжение)</i>										
	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20
94	4	A124	20	A143	A151	2	A173	1	A191	A201
95	4	A121	54	A143	A152	1	A173	1	A192	A201
96	4	A124	58	A143	A151	1	A173	1	A192	A201
97	2	A123	61	A143	A152	1	A173	1	A191	A201
98	4	A123	34	A143	A152	1	A173	1	A191	A201
99	4	A121	36	A143	A152	1	A173	1	A191	A201
100	4	A122	36	A141	A151	2	A174	2	A192	A201

## **A.2 Описание исходных данных**

Описание всех 20 характеристик физических лиц из используемой базы кредитных историй представлено следующим образом.

### 1. Баланс счёта до востребования:

A11 – отрицательное значение счёта;

A12 – от 0 до 100 тыс. рублей;

A13 – более 100 тыс. рублей ;

A14 – отсутствие счёта.

### 2. Продолжительность кредита, измеряется в месяцах.

### 3. Кредитная история в данном банке:

A30 – кредит берётся в первый раз;

A31 – все кредиты в данном банке выплачены в срок;

A32 – действующие кредиты в данном банке выплачиваются в срок;

A33 – отсрочка платежа во взятых ранее кредитах;

A34 – рискованный заёмщик.

### 4. Цель взятия кредита:

A40 – новый автомобиль;

A41 – подержанный автомобиль;

A42 – мебель;

A43 – телевизор или акустическое оборудование;

A44 – бытовая техника;

A45 – ремонт;

A46 – образование;

A47 – отпуск;

A48 – переквалификация;

A49 – работа;

A410 – другое.

5. Сумма кредита.

6. Сбережения, в том числе и облигации:

A61 – до 50 тыс. рублей;

A62 – от 50 до 250 тыс. рублей;

A63 – от 250 до 500 тыс. рублей;

A64 – более 500 тыс. рублей;

A65 – отсутствуют данные или отсутствуют сбережения.

7. Продолжительность трудоустройства на последнем рабочем месте:

A71 – нетрудоустроен;

A72 – до 1 года;

A73 – от 1 года до 4 лет;

A74 – от 4 до 7 лет;

A75 – более 7 лет.

8. Доля платежей в располагаемом доходе.

9. Пол и семейный статус:

A91 – мужской, разведён;

A92 – женский, замужем;

A93 – мужской, не женат;

A94 – мужской, женат;

A95 – женский, не замужем.

10. Созаявитель на кредит или поручитель:

A101 – отсутствует;

A102 – есть созаявитель на кредит;

A103 – есть поручитель.

11. Продолжительность проживания в нынешних апартаментах.

12. Собственность:

A121 – жилая площадь;

A122 – отсутствует жилищная площадь, но есть счёт в строительной сберегательной кассе;

A123 – отсутствуют первые два пункта, но есть автомобиль;

A124 – отсутствует информация или нет собственности.

13. Возраст, измеряется в годах.

14. Действующие кредиты в других местах:

A141 – в другом банке;

A142 – в магазинах;

A143 – отсутствует.

15. Плата за жилую площадь:

A151 – арендная плата;

A152 – в собственности (коммунальные платежи);

A153 – отсутствуют платежи.

16. Число действующих кредитов в данном банке.

17. Вид трудоустройства:

A171 – безработный;

A172 – работник без квалификации;

A173 – квалифицированный работник;

A174 – управляющий или работник высшей квалификации.

18. Количество людей, готовых предоставить средства на существование.

19. Наличие городского телефона:

A191 – телефон отсутствует;

A192 – в наличии зарегистрированный городской номер на имя заявителя.

20. Иностранец:

A201 – да;

A202 – нет.

### А.3 Нормированные значения качественных характеристик

Одним из этапов вычисления оценки кредитоспособности является нормирование исходных характеристик в промежуток от нуля до единицы ( $[0,1]$ ) посредством использования функций, оценивающих качество с точки зрения кредитоспособности по отдельно взятой характеристике. Сверху таблицы представлена структура весовых коэффициентов.

	$w_1$		$w_3$					
	$w_{11}$	$w_{12}$	$w_{31}$		$w_{32}$			
			$w_{311}$	$w_{312}$	$w_{321}$	$w_{322}$	$w_{323}$	$w_{324}$
	A9	A13	A1	A6	A11	A12	A15	A19
1	0,2	0,000	0,0	0,00	1,00	1,0	1,0	1
2	0,8	0,028	0,5	0,05	0,50	1,0	1,0	0
3	0,2	0,575	0,0	0,05	0,75	1,0	1,0	0
4	0,2	0,844	0,0	0,05	1,00	0,6	0,7	0
<i>продолжение следует</i>								

<i>(продолжение)</i>								
	A9	A13	A1	A6	A11	A12	A15	A19
5	0,2	0,282	0,0	0,05	1,00	0,0	0,7	0
6	0,2	0,844	0,0	0,00	1,00	0,0	0,7	1
7	0,2	0,282	0,0	0,75	1,00	0,6	1,0	0
8	0,2	0,844	0,5	0,05	0,50	0,4	0,0	1
9	0,6	0,000	0,0	1,00	1,00	1,0	1,0	0
10	1,0	0,352	0,5	0,05	0,50	0,4	1,0	0
11	0,8	0,156	0,5	0,05	0,25	0,4	0,0	0
12	0,8	0,104	0,0	0,05	1,00	0,6	0,0	0
13	0,8	0,028	0,5	0,05	0,25	0,4	1,0	1
14	0,2	0,000	0,0	0,05	1,00	0,4	1,0	0
15	0,8	0,352	0,0	0,05	1,00	0,4	0,0	0
16	0,8	0,648	0,0	0,30	0,50	0,4	1,0	0
17	0,2	0,282	0,0	0,00	1,00	0,6	1,0	0
18	0,2	0,156	0,0	0,00	0,75	0,4	1,0	0
19	0,8	0,896	0,5	0,05	0,50	0,0	0,7	1
20	0,2	0,575	0,0	0,75	0,50	0,4	1,0	1
21	0,2	0,648	0,0	0,05	1,00	0,4	1,0	1
22	0,2	0,896	0,0	0,75	0,75	1,0	0,0	0
23	0,2	0,648	0,0	0,05	0,75	1,0	0,0	0
24	0,2	0,896	0,5	0,30	1,00	0,6	1,0	0
25	1,0	0,216	0,0	0,00	0,25	0,4	1,0	0
26	0,2	0,896	0,0	0,05	0,50	1,0	1,0	1
27	1,0	0,993	0,0	0,05	1,00	0,4	1,0	0
28	0,8	0,972	1,0	1,00	0,75	1,0	0,0	0
29	0,2	0,784	0,5	0,05	0,50	1,0	1,0	0
30	0,2	0,000	0,0	0,05	1,00	0,0	1,0	1
31	1,0	0,896	0,5	1,00	0,75	1,0	1,0	1
32	0,2	0,282	0,0	0,05	0,50	0,4	1,0	0
33	0,2	0,500	0,5	0,30	0,50	0,4	1,0	1
34	0,2	0,061	0,0	0,00	1,00	0,0	0,0	0
35	0,8	0,718	1,0	0,05	0,25	0,6	1,0	1
36	0,2	0,156	0,5	0,05	0,50	0,6	1,0	0
37	0,2	0,575	0,0	0,05	0,75	0,0	0,7	1
38	0,2	0,939	1,0	0,05	0,50	1,0	1,0	0
39	0,2	0,939	1,0	0,05	0,50	0,4	1,0	1
40	0,2	0,104	0,5	0,05	0,75	1,0	1,0	0
41	0,2	0,500	0,0	0,75	0,50	0,4	1,0	0
42	0,6	0,216	0,5	0,75	0,25	0,4	1,0	1
43	0,2	0,896	0,5	0,05	1,00	1,0	1,0	1
44	1,0	0,104	0,0	0,30	1,00	0,4	0,0	0
45	0,8	0,028	0,0	0,05	1,00	0,0	0,7	0
46	0,8	0,844	0,0	0,05	1,00	0,4	1,0	0
47	0,2	0,993	0,0	0,75	1,00	0,4	1,0	0
48	0,8	0,061	0,0	0,75	0,50	0,6	0,0	1

*продолжение следует*

<i>(продолжение)</i>								
	A9	A13	A1	A6	A11	A12	A15	A19
49	0,2	0,993	0,0	0,05	1,00	0,6	1,0	0
50	0,8	0,352	0,0	0,30	0,50	1,0	1,0	0
51	0,2	0,425	0,5	0,00	0,50	0,6	1,0	0
52	0,2	0,500	0,5	0,05	0,50	0,4	1,0	1
53	0,2	0,156	0,0	0,05	0,50	0,4	1,0	0
54	0,2	0,575	0,0	0,00	0,25	0,6	1,0	1
55	0,2	0,061	0,5	0,05	1,00	0,0	0,7	1
56	0,2	0,216	0,0	0,00	0,50	1,0	1,0	0
57	0,2	0,352	0,5	0,00	0,25	0,0	1,0	1
58	0,8	0,575	0,0	0,05	0,50	0,4	1,0	0
59	0,8	0,061	1,0	0,05	0,50	0,4	1,0	0
60	0,8	0,061	0,0	0,05	1,00	0,0	0,0	1
61	1,0	0,282	0,5	0,05	0,25	1,0	1,0	1
62	0,2	0,500	0,5	0,00	1,00	1,0	1,0	1
63	0,2	0,000	0,5	0,05	1,00	0,0	0,7	1
64	0,2	0,156	0,5	0,05	0,50	0,4	1,0	1
65	0,8	0,216	0,0	0,05	1,00	0,6	1,0	1
66	0,2	0,648	0,0	0,00	1,00	0,6	1,0	1
67	0,8	0,425	0,0	0,05	0,50	0,4	1,0	0
68	1,0	0,028	0,5	1,00	0,25	1,0	1,0	0
69	0,2	0,939	0,0	0,05	1,00	0,0	0,7	1
70	0,8	0,156	0,0	0,00	1,00	0,4	1,0	0
71	0,8	0,500	0,0	0,05	0,50	0,6	1,0	0
72	0,2	0,784	0,0	0,00	0,50	0,6	0,0	1
73	0,2	0,425	0,0	0,05	1,00	0,0	0,7	1
74	0,8	0,993	0,5	0,05	0,25	1,0	1,0	0
75	0,2	1,000	0,0	0,00	1,00	0,0	1,0	1
76	0,2	0,000	0,0	0,05	1,00	0,0	0,7	0
77	0,2	0,784	0,0	0,05	0,75	0,4	1,0	0
78	0,2	0,425	0,5	0,05	1,00	0,6	1,0	0
79	0,2	0,993	0,0	0,00	0,50	0,6	1,0	0
80	1,0	0,028	0,5	0,05	0,25	0,6	1,0	0
81	0,8	0,896	0,0	0,00	0,25	0,4	1,0	1
82	0,2	0,718	0,0	0,75	0,75	0,6	1,0	1
83	0,8	0,104	0,0	0,30	1,00	0,6	0,0	0
84	0,8	0,028	0,0	0,05	1,00	1,0	1,0	1
85	0,2	0,352	0,0	0,05	1,00	1,0	1,0	0
86	0,8	0,425	0,0	0,05	0,50	1,0	1,0	1
87	0,8	0,282	0,5	0,05	0,25	0,6	1,0	0
88	0,2	0,718	0,5	0,30	1,00	0,0	0,7	1
89	0,2	0,500	0,0	0,30	0,75	0,4	1,0	1
90	0,2	0,352	0,0	0,05	0,75	1,0	1,0	0
91	0,2	0,104	0,0	0,05	1,00	1,0	1,0	0
92	0,2	0,216	0,0	0,05	0,75	1,0	1,0	0

*продолжение следует*

<i>(продолжение)</i>								
	A9	A13	A1	A6	A11	A12	A15	A19
93	0,8	0,718	0,0	0,00	0,75	0,6	1,0	0
94	0,2	0,000	1,0	0,00	1,00	0,0	0,0	0
95	0,2	0,216	0,5	1,00	1,00	1,0	1,0	1
96	0,2	0,028	0,5	0,05	1,00	0,0	0,0	1
97	0,8	0,000	0,0	0,00	0,50	0,4	1,0	0
98	0,2	0,784	0,5	0,30	1,00	0,4	1,0	0
99	0,2	0,896	0,5	0,05	1,00	1,0	1,0	0
100	0,2	0,896	0,5	0,00	1,00	0,6	0,0	1

	$w_2$				$w_4$				
	$w_{21}$	$w_{22}$	$w_{23}$	$w_{24}$	$w_{41}$			$w_{42}$	
					$w_{411}$	$w_{412}$	$w_{413}$	$w_{421}$	$w_{422}$
	A7	A17	A8	A20	A3	A14	A16	A10	A18
1	1,00	0,6	1,00	0	0,0	1,00	0,50	0,0	0,5
2	0,36	0,6	0,50	0	0,6	1,00	0,25	0,0	0,5
3	0,79	0,2	0,50	0	0,0	1,00	0,25	0,0	1,0
4	0,79	0,6	0,50	0	0,6	1,00	0,25	1,0	1,0
5	0,36	0,6	0,75	0	0,2	1,00	0,50	0,0	1,0
6	0,36	0,2	0,50	0	0,6	1,00	0,25	0,0	1,0
7	1,00	0,6	0,75	0	0,6	1,00	0,25	0,0	0,5
8	0,36	1,0	0,50	0	0,6	1,00	0,25	0,0	0,5
9	0,79	0,2	0,50	0	0,6	1,00	0,25	0,0	0,5
10	0,00	1,0	1,00	0	0,0	1,00	0,50	0,0	0,5
11	0,14	0,6	0,75	0	0,6	1,00	0,25	0,0	0,5
12	0,14	0,6	0,75	0	0,6	1,00	0,25	0,0	0,5
13	0,36	0,6	0,25	0	0,6	1,00	0,25	0,0	0,5
14	1,00	0,2	1,00	0	0,0	1,00	0,50	0,0	0,5
15	0,36	0,6	0,50	0	0,6	1,00	0,25	0,0	0,5
16	0,36	0,2	1,00	0	0,6	1,00	0,25	0,0	0,5
17	1,00	0,6	1,00	0	0,0	1,00	0,50	0,0	0,5
18	0,14	0,6	0,50	0	0,0	0,33	0,75	0,0	0,5
19	1,00	1,0	1,00	0	0,6	1,00	0,25	0,0	0,5
20	1,00	0,6	0,75	0	0,6	1,00	0,25	0,0	1,0
21	0,36	0,6	1,00	0	0,0	1,00	0,75	0,0	0,5
22	0,36	0,6	0,50	0	0,6	1,00	0,25	0,0	1,0
23	0,14	0,2	0,25	1	0,0	1,00	0,50	0,0	1,0
24	0,14	0,6	0,75	0	0,0	1,00	0,25	0,0	0,5
25	0,36	0,6	0,50	1	0,0	1,00	0,50	0,0	0,5
26	0,36	0,2	0,25	0	0,6	0,33	0,25	0,0	0,5
27	1,00	0,2	1,00	0	0,0	1,00	0,25	0,0	0,5
28	0,36	0,6	0,75	0	1,0	1,00	0,50	0,0	0,5
29	0,36	0,6	0,75	0	0,6	1,00	0,25	1,0	0,5
30	1,00	0,6	0,75	0	0,2	1,00	0,50	0,0	0,5

*продолжение следует*

<i>(продолжение)</i>									
	A7	A17	A8	A20	A3	A14	A16	A10	A18
31	0,14	0,6	0,75	0	0,6	0,33	0,25	0,0	0,5
32	0,36	0,6	0,50	0	0,6	0,67	0,25	0,0	0,5
33	0,36	0,6	0,50	0	0,6	1,00	0,50	0,0	0,5
34	1,00	0,2	1,00	0	0,0	1,00	0,25	0,0	0,5
35	0,14	1,0	1,00	0	0,6	0,33	0,25	0,0	0,5
36	0,14	0,2	1,00	0	0,0	1,00	0,50	0,0	0,5
37	0,36	0,6	0,25	0	0,0	0,33	0,25	0,0	0,5
38	0,36	0,6	1,00	0	0,6	0,67	0,25	0,8	0,5
39	0,36	0,6	0,50	0	0,6	1,00	0,25	0,0	0,5
40	0,36	0,6	1,00	0	0,6	1,00	0,25	0,0	0,5
41	1,00	1,0	1,00	0	0,6	0,33	0,25	0,0	0,5
42	0,36	0,6	0,75	0	0,6	1,00	0,25	0,0	0,5
43	0,36	0,2	0,50	0	0,2	1,00	0,25	0,0	1,0
44	0,79	0,6	0,25	0	0,0	1,00	0,50	0,0	0,5
45	1,00	0,2	1,00	0	0,0	0,67	0,50	0,0	0,5
46	0,14	1,0	1,00	0	0,0	1,00	0,50	0,0	0,5
47	1,00	0,6	1,00	0	0,6	1,00	0,25	0,0	0,5
48	0,00	0,0	0,25	0	0,6	1,00	0,25	0,0	0,5
49	0,36	0,2	0,25	0	0,0	1,00	0,50	0,0	0,5
50	0,36	0,6	1,00	0	0,6	1,00	0,25	0,8	0,5
51	0,14	0,2	1,00	0	0,2	0,33	0,25	0,0	0,5
52	1,00	1,0	0,25	0	0,2	1,00	0,50	0,0	0,5
53	0,36	0,6	0,75	0	0,6	1,00	0,25	0,0	0,5
54	0,36	0,6	0,50	0	0,6	1,00	0,25	0,0	0,5
55	1,00	0,6	1,00	0	0,2	0,33	0,50	0,0	0,5
56	0,36	0,2	0,25	0	1,0	0,67	0,25	1,0	1,0
57	0,00	1,0	0,50	0	0,6	1,00	0,25	0,0	0,5
58	0,36	0,6	0,50	0	0,0	0,67	0,50	0,0	0,5
59	1,00	1,0	0,75	0	0,6	1,00	0,25	0,0	0,5
60	0,14	0,2	1,00	0	0,0	1,00	0,50	0,8	0,5
61	0,36	0,6	0,50	0	0,6	0,33	0,25	0,0	0,5
62	1,00	0,6	1,00	0	0,0	1,00	0,50	1,0	0,5
63	1,00	1,0	1,00	0	0,0	1,00	0,25	0,0	0,5
64	0,36	0,6	0,50	0	0,0	1,00	0,25	0,0	0,5
65	0,14	0,6	1,00	0	0,6	1,00	0,25	0,0	0,5
66	1,00	0,6	1,00	0	0,6	1,00	1,00	0,0	1,0
67	0,14	0,6	0,50	0	0,6	0,33	0,25	0,0	0,5
68	0,36	0,6	1,00	0	0,6	1,00	0,25	0,0	0,5
69	0,36	0,6	1,00	0	0,6	0,67	0,25	0,0	0,5
70	0,36	0,6	1,00	0	0,6	1,00	0,25	0,0	0,5
71	0,36	0,6	0,25	0	0,6	0,33	0,25	0,0	0,5
72	1,00	0,2	1,00	0	0,0	1,00	0,50	0,0	0,5
73	1,00	1,0	0,75	0	0,0	0,33	0,50	0,0	1,0
74	0,79	0,2	0,50	0	0,0	0,33	0,50	0,0	0,5

*продолжение следует*



<i>(продолжение)</i>									
	A7	A17	A8	A20	A3	A14	A16	A10	A18
75	1,00	1,0	1,00	0	0,6	1,00	0,25	0,0	0,5
76	1,00	1,0	1,00	0	0,0	1,00	0,50	0,0	0,5
77	0,14	0,6	1,00	0	0,6	1,00	0,25	0,0	0,5
78	0,79	0,6	0,50	0	0,2	1,00	0,25	0,0	0,5
79	0,36	0,2	0,50	0	0,0	1,00	0,25	0,0	1,0
80	0,14	0,6	0,50	0	0,6	1,00	0,25	0,0	0,5
81	0,14	0,6	0,25	0	0,6	1,00	0,50	0,0	0,5
82	1,00	0,6	1,00	0	0,6	0,67	0,25	0,0	0,5
83	0,36	0,2	0,75	0	0,6	1,00	0,25	0,0	0,5
84	1,00	0,2	1,00	0	0,6	1,00	0,25	1,0	0,5
85	1,00	0,2	0,75	0	0,6	1,00	0,25	0,0	0,5
86	0,36	1,0	1,00	0	0,0	1,00	0,50	1,0	0,5
87	0,14	0,6	1,00	0	0,0	1,00	0,50	0,0	0,5
88	0,36	0,6	0,25	0	0,6	1,00	0,25	0,0	1,0
89	0,79	1,0	1,00	0	0,6	1,00	0,25	0,0	1,0
90	0,79	0,6	1,00	0	0,0	1,00	0,50	0,0	0,5
91	1,00	0,6	1,00	0	0,0	1,00	0,25	0,0	0,5
92	1,00	0,6	1,00	0	0,0	1,00	0,25	0,0	0,5
93	1,00	0,2	1,00	0	0,0	0,33	0,25	0,0	1,0
94	1,00	0,6	1,00	0	0,0	1,00	0,50	0,8	0,5
95	1,00	0,6	1,00	0	0,6	1,00	0,25	0,0	0,5
96	0,14	0,6	0,75	0	0,0	1,00	0,25	0,0	0,5
97	0,79	0,6	1,00	0	0,0	1,00	0,25	0,0	0,5
98	0,36	0,6	1,00	0	0,6	1,00	0,25	0,0	0,5
99	1,00	0,6	1,00	0	0,0	1,00	0,25	0,0	0,5
100	0,79	1,0	0,75	0	0,2	0,33	0,50	0,0	1,0

## А.4 Программный код

В данном разделе приведена часть кода используемая в процессе вычисления класса кредитоспособности для программной среды Wolfram Mathematica 8. Для вычисления весовых коэффициентов третьего и четвёртого уровней иерархии сначала действует условный цикл, перебирающий комбинации весовых коэффициентов. Из комбинаций, удовлетворяющих нормирующему условию, и дополнительной информации составляются списки допустимых весовых коэффициентов. Интервальная информация задана в виде границ действующего цикла. Списки допустимых комбинаций весовых коэффициентов для вычисления показателей 3-го уровня иерархии составляются с использованием следующего кода:

$W411 = 0; W412 = 0; W413 = 0; ListW411 = \{\}; ListW412 = \{\};$   
 $ListW413 = \{\}; Do[If[W411 + W412 + W413 == 1,$   
 $If[W412 > W411, If[W411 > W413,$   
 $ListW411 = Prepend[ListW411, W411];$   
 $ListW412 = Prepend[ListW412, W412];$   
 $ListW413 = Prepend[ListW413, W413], 0], 0], 0],$   
 $\{W411, 0.33, 1, 0.01\}, \{W412, 0, 1, 0.01\}, \{W413, 0, 1, 0.01\}]$

$W421 = 0; W422 = 0; ListW421 = \{\}; ListW422 = \{\};$   
 $Do[If[W421 + W422 == 1, If[W421 < W422,$   
 $ListW421 = Prepend[ListW421, W421];$   
 $ListW422 = Prepend[ListW422, W422], 0], 0],$   
 $\{W421, 0, 0.2, 0.01\}, \{W422, 0.1, 1, 0.01\}]$

$W321 = 0; W322 = 0; W323 = 0; W324 = 0; ListW321 = \{\};$   
 $ListW322 = \{\}; ListW323 = \{\}; ListW324 = \{\};$   
 $Do[If[W321 + W322 + W323 + W324 == 1, If[W321 > W323,$   
 $If[W322 > W321, ListW321 = Prepend[ListW321, W321];$   
 $ListW322 = Prepend[ListW322, W322];$   
 $ListW323 = Prepend[ListW323, W323];$   
 $ListW324 = Prepend[ListW324, W324], 0], 0], 0],$   
 $\{W321, 0, 1, 0.01\}, \{W322, 0, 1, 0.01\},$   
 $\{W323, 0, 1, 0.01\}, \{W324, 0, 0.1, 0.01\}]$

```

W311 = 0; W312 = 0; ListW311 = {}; ListW312 = {};
Do[If[W311 + W312 == 1, If[W312 > W311,
ListW311 = Prepend[ListW311, W311];
ListW312 = Prepend[ListW312, W312], 0], 0],
{W311, 0.2, 1, 0.01}, {W312, 0, 1, 0.01}]

```

Списки допустимых комбинаций весовых коэффициентов для вычисления показателей 2-го уровня иерархии составляются с использованием следующего кода:

```

W21 = 0; W22 = 0; W23 = 0; W24 = 0; ListW21 = {}; ListW22 = {};
ListW23 = {}; ListW24 = {}; Do[If[W21 + W22 + W23 + W24 == 1,
If[W22 > W23, If[W23 > W21, ListW21 = Prepend[ListW21, W21];
ListW22 = Prepend[ListW22, W22];
ListW23 = Prepend[ListW23, W23];
ListW24 = Prepend[ListW24, W24], 0], 0], 0], {W21, 0.1, 1, 0.01},
{W22, 0, 1, 0.01}, {W23, 0.2, 1, 0.01}, {W24, 0, 0.2, 0.01}]

```

```

W11 = 0; W12 = 0; ListW11 = {}; ListW12 = {}; Do[If[W11 + W12 == 1,
If[W11 > W12, ListW11 = Prepend[ListW11, W11];
ListW12 = Prepend[ListW12, W12], 0], 0],
{W11, 0, 1, 0.01}, {W12, 0, 1, 0.01}]

```

```

W41 = 0; W42 = 0; ListW41 = {}; ListW42 = {};
Do[If[W41 + W42 == 1, If[W41 > W42, ListW1 = Prepend[ListW1, W1];
ListW2 = Prepend[ListW2, W2], 0], 0],
{W41, 0, 1, 0.01}, {W42, 0.3, 1, 0.01}]

```

Списки допустимых комбинаций весовых коэффициентов второго уровня иерархии составляются посредством реализации следующего кода:

```

W1 = 0; W2 = 0; W3 = 0; W4 = 0; ListW1 = {}; ListW2 = {};
ListW3 = {}; ListW4 = {}; Do[If[W1 + W2 + W3 + W4 == 1,
If[W2 > W4, If[W2 > W3, If[W3 > W1, ListW1 = Prepend[ListW1, W1];
ListW2 = Prepend[ListW2, W2]; ListW3 = Prepend[ListW3, W3];
ListW4 = Prepend[ListW4, W4], 0], 0], 0], 0],
{W1, 0.05, 1, 0.01}, {W2, 0, 1, 0.01},
{W3, 0, 1, 0.01}, {W4, 0, 1, 0.01}]

```

Далее вычисляются среднее значение и стандартное отклонение всех весовых коэффициентов, например,  $W311$ :

$$MeanW311 = Mean[ListW311]$$

Сначала необходимо вычислить агрегированные оценки характеристик второго уровня. Задаются пустые списки, которые затем заполняются вычислениями показателей третьего уровня иерархии, соответствующие весовым коэффициентам  $w_{42}$ ,  $w_{41}$ ,  $w_{32}$ ,  $w_{31}$ :

```

Do[Level1H[[i,j]] = Table[0, {i, 1, 1000}, {i, j, 4};
Level2H[[i,j]] = Table[0, {i, 1, 1000}, {i, j, 4};

```

$$Do[Level2H[[i,4]] = Mean[ListW421 * NH[i][[10]] + ListW422 * NH[i][[18]]], \{i, 1, 1000\}];$$

$$Do[Level2H[[i,3]] = Mean[ListW411 * NH[i][[3]] + ListW412 * NH[i][[14]] + ListW412 * NH[i][[16]]], \{i, 1, 1000\}];$$

$$Do[Level2H[[i,2]] = Mean[ListW321 * NH[i][[11]] + ListW322 * NH[i][[12]] + ListW323 * NH[i][[15]] + ListW324 * NH[i][[19]]], \{i, 1, 1000\}];$$

$$Do[Level2H[[i,1]] = Mean[ListW311 * NH[i][[1]] + ListW312 * NH[i][[6]]], \{i, 1, 1000\}];$$

Для вычисления показателей второго уровня иерархии, соответствующие весовым коэффициентам  $w_4$ ,  $w_3$ ,  $w_2$ ,  $w_1$  составлен следующий код, который необходимо реализовывать в приведенной последовательности:

$$Do[Level1H[[i,4]] = Mean[ListW41 * Level2H[i][[3]] + ListW42 * Level2H[i][[4]]], \{i, 1, 1000\}];$$

$$Do[Level1H[[i,3]] = Mean[ListW31 * Level2H[i][[1]] + ListW32 * Level2H[i][[2]]], \{i, 1, 1000\}];$$

$$Do[Level1H[[i,2]] = Mean[ListW21 * NH[i][[7]] + ListW22 * NH[i][i][[17]] + ListW23 * NH[i][[8]] + ListW24 * NH[i][i][[20]]], \{i, 1, 1000\}];$$

$$Do[Level1H[[i,1]] = Mean[ListW11 * Level2H[i][[9]] + ListW12 * Level2H[i][[13]]], \{i, 1, 1000\}];$$

## А.5 Значение класса кредитоспособности

В нижеприведённой таблице представлены значения класса кредитоспособности физических лиц из используемой базы кредитных историй, которые построены по описанному в работе методу для тестового множества.

№	Оценка	№	Оценка	№	Оценка	№	Оценка	№	Оценка
1	0,56121	101	0,52746	201	0,50806	301	0,56561	401	0,47558
2	0,54698	102	0,59218	202	0,53826	302	0,47595	402	0,46356
3	0,58576	103	0,49619	203	0,63657	303	0,49901	403	0,47324
4	0,47402	104	0,67592	204	0,55709	304	0,56544	404	0,60942
5	0,56119	105	0,57523	205	0,61668	305	0,31667	405	0,70714
6	0,49627	106	0,56949	206	0,55660	306	0,49244	406	0,59889
7	0,47656	107	0,56067	207	0,54979	307	0,51604	407	0,49782
8	0,59616	108	0,61518	208	0,53232	308	0,56836	408	0,53067
9	0,59870	109	0,53745	209	0,65809	309	0,45005	409	0,34030
10	0,68633	110	0,54831	210	0,43980	310	0,32787	410	0,41342
11	0,58174	111	0,55619	211	0,62620	311	0,56698	411	0,58950
12	0,55639	112	0,47837	212	0,66210	312	0,51856	412	0,30692
13	0,52371	113	0,53337	213	0,69909	313	0,53406	413	0,57134
14	0,62566	114	0,53011	214	0,42013	314	0,49222	414	0,56489
15	0,50284	115	0,59005	215	0,66085	315	0,59052	415	0,56770
16	0,63817	116	0,45991	216	0,43967	316	0,47945	416	0,53924
17	0,41866	117	0,52567	217	0,55558	317	0,49543	417	0,54011
18	0,45862	118	0,46057	218	0,62391	318	0,46643	418	0,50393
19	0,56591	119	0,55751	219	0,48576	319	0,60383	419	0,45117
20	0,55056	120	0,47300	220	0,54674	320	0,55082	420	0,68877
21	0,50537	121	0,48024	221	0,60441	321	0,58235	421	0,53017
22	0,52128	122	0,41560	222	0,59810	322	0,56648	422	0,66623
23	0,47943	123	0,52547	223	0,49300	323	0,50147	423	0,52416

продолжение следует

<i>(продолжение)</i>									
№	Оценка	№	Оценка	№	Оценка	№	Оценка	№	Оценка
24	0,58558	124	0,54213	224	0,52132	324	0,46928	424	0,59166
25	0,59822	125	0,51764	225	0,54841	325	0,58523	425	0,43706
26	0,46758	126	0,59465	226	0,36929	326	0,61840	426	0,48325
27	0,55339	127	0,50380	227	0,41432	327	0,46805	427	0,51556
28	0,41885	128	0,53706	228	0,51966	328	0,44535	428	0,58434
29	0,56910	129	0,55595	229	0,63939	329	0,53465	429	0,57915
30	0,44996	130	0,39729	230	0,54407	330	0,55084	430	0,46664
31	0,58669	131	0,51683	231	0,40335	331	0,56852	431	0,49950
32	0,53590	132	0,49997	232	0,41314	332	0,55278	432	0,56150
33	0,58648	133	0,51636	233	0,59030	333	0,48472	433	0,48042
34	0,64429	134	0,51991	234	0,54529	334	0,57863	434	0,62312
35	0,51335	135	0,44346	235	0,52945	335	0,43023	435	0,50660
36	0,45401	136	0,57801	236	0,34721	336	0,43377	436	0,57036
37	0,64129	137	0,63232	237	0,57920	337	0,56466	437	0,48303
38	0,58733	138	0,47522	238	0,46779	338	0,40720	438	0,68373
39	0,61436	139	0,60967	239	0,55146	339	0,55084	439	0,53442
40	0,34830	140	0,49303	240	0,46123	340	0,66415	440	0,56788
41	0,48383	141	0,41759	241	0,48541	341	0,57490	441	0,58458
42	0,54530	142	0,51573	242	0,45863	342	0,36841	442	0,50561
43	0,60059	143	0,61419	243	0,59641	343	0,57633	443	0,54323
44	0,44675	144	0,65298	244	0,56949	344	0,53084	444	0,42376
45	0,46438	145	0,69927	245	0,64389	345	0,57801	445	0,60170
46	0,53182	146	0,49343	246	0,29251	346	0,52401	446	0,45470
47	0,55417	147	0,53386	247	0,44076	347	0,45669	447	0,60037
48	0,51350	148	0,53574	248	0,45319	348	0,59299	448	0,60129
49	0,46711	149	0,55788	249	0,42956	349	0,54480	449	0,55714
50	0,45971	150	0,47775	250	0,50806	350	0,47037	450	0,54195
51	0,59130	151	0,63598	251	0,49651	351	0,60073	451	0,40924
52	0,45583	152	0,45929	252	0,35505	352	0,50943	452	0,47020
53	0,60836	153	0,63459	253	0,47721	353	0,50380	453	0,55965
54	0,48613	154	0,62653	254	0,57198	354	0,53985	454	0,64846
55	0,72385	155	0,54872	255	0,61030	355	0,52887	455	0,64467
56	0,41316	156	0,38562	256	0,60743	356	0,61504	456	0,71801
57	0,51082	157	0,48255	257	0,35526	357	0,62590	457	0,44006
58	0,39586	158	0,54981	258	0,47912	358	0,48085	458	0,43164
59	0,50612	159	0,47971	259	0,54024	359	0,55387	459	0,39797
60	0,34698	160	0,47422	260	0,56964	360	0,57930	460	0,59829
61	0,54417	161	0,58733	261	0,48245	361	0,54369	461	0,55775
62	0,48457	162	0,58302	262	0,55455	362	0,56807	462	0,45205
63	0,42741	163	0,59430	263	0,43927	363	0,59833	463	0,48693
64	0,44207	164	0,57207	264	0,63731	364	0,44747	464	0,53310
65	0,56975	165	0,39035	265	0,61011	365	0,31712	465	0,44721
66	0,56857	166	0,53302	266	0,54418	366	0,52164	466	0,57814
67	0,54320	167	0,61797	267	0,39925	367	0,52552	467	0,43006

*продолжение следует*

<i>(продолжение)</i>									
№	Оценка	№	Оценка	№	Оценка	№	Оценка	№	Оценка
68	0,51752	168	0,48645	268	0,48145	368	0,45926	468	0,65058
69	0,61631	169	0,56382	269	0,46526	369	0,56840	469	0,57012
70	0,55127	170	0,56568	270	0,39926	370	0,63106	470	0,35669
71	0,47789	171	0,38518	271	0,57691	371	0,55815	471	0,60747
72	0,53979	172	0,52028	272	0,58971	372	0,56293	472	0,38249
73	0,58184	173	0,62801	273	0,55746	373	0,55320	473	0,36753
74	0,48461	174	0,46722	274	0,49698	374	0,56641	474	0,62483
75	0,50756	175	0,41082	275	0,22202	375	0,45325	475	0,54756
76	0,48683	176	0,66847	276	0,59037	376	0,59153	476	0,47859
77	0,56648	177	0,56130	277	0,54876	377	0,59043	477	0,57711
78	0,51708	178	0,53559	278	0,67253	378	0,58485	478	0,57546
79	0,52297	179	0,43551	279	0,68487	379	0,56454	479	0,50154
80	0,48617	180	0,53756	280	0,48567	380	0,48141	480	0,56521
81	0,53885	181	0,43421	281	0,53583	381	0,64168	481	0,46222
82	0,27683	182	0,53472	282	0,58878	382	0,47944	482	0,51123
83	0,58702	183	0,54001	283	0,36191	383	0,59003	483	0,53713
84	0,41589	184	0,49500	284	0,41934	384	0,36265	484	0,50380
85	0,70700	185	0,41766	285	0,54858	385	0,57654	485	0,37741
86	0,42291	186	0,57263	286	0,49606	386	0,55970	486	0,54114
87	0,48774	187	0,55383	287	0,58390	387	0,57862	487	0,46983
88	0,47819	188	0,56783	288	0,59591	388	0,46777	488	0,50045
89	0,51217	189	0,62006	289	0,54662	389	0,66364	489	0,58990
90	0,51138	190	0,51263	290	0,42438	390	0,56355	490	0,45793
91	0,50114	191	0,46357	291	0,62493	391	0,66848	491	0,36123
92	0,56644	192	0,47769	292	0,69218	392	0,55150	492	0,51743
93	0,53983	193	0,46261	293	0,45834	393	0,40412	493	0,46117
94	0,45357	194	0,33438	294	0,57884	394	0,49028	494	0,61396
95	0,68530	195	0,56980	295	0,60323	395	0,50817	495	0,57585
96	0,38644	196	0,48221	296	0,49859	396	0,61441	496	0,51226
97	0,61102	197	0,69786	297	0,51852	397	0,66591	497	0,67657
98	0,34948	198	0,27533	298	0,35879	398	0,63058	498	0,59800
99	0,41518	199	0,49008	299	0,61731	399	0,38168	499	0,52002
100	0,39729	200	0,70513	300	0,40731	400	0,50856	500	0,45063