

ISSN 2311-8709 (Online)  
ISSN 2071-4688 (Print)



ВЫХОДИТ 4 РАЗА В МЕСЯЦ

# ФИНАНСЫ<sup>®</sup> & КРЕДИТ

НАУЧНО-ПРАКТИЧЕСКИЙ И ТЕОРЕТИЧЕСКИЙ ЖУРНАЛ

**2015** выпуск 27  
ИЮЛЬ



# FINANCE & CREDIT

A peer reviewed analytical and practical journal  
2015, July  
Issue 27



Научно-практический и теоретический журнал

Основан в 1994 году  
Журнал выходит 4 раза в месяц  
Статьи рецензируются

Журнал рекомендован ВАК Минобрнауки России для публикации научных работ, отражающих основное научное содержание кандидатских и докторских диссертаций  
Журнал реферируется в ВИНТИ РАН  
Журнал включен в Российский индекс научного цитирования (РИНЦ)  
Журнал зарегистрирован в Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций.  
Свидетельство ПИ № ФС77-60938 от 02 марта 2015 г.  
Предыдущее Свидетельство о регистрации № 013007 от 03 ноября 1994 г. выдано Комитетом Российской Федерации по печати

**Учредитель:**

ООО «Издательский дом ФИНАНСЫ и КРЕДИТ»  
Юр. адрес: 111141, г. Москва, Зелёный проспект, д.8, кв. 1  
Факт. адрес: 111397, г. Москва, Зелёный проспект, д. 20  
Почтовый адрес: 111401, г. Москва, а/я 10

**Издатель:**

ООО «Информационный центр «Финансы и кредит»  
Юр. адрес: 123182, г. Москва, ул. Авиационная, 79-480  
Факт. адрес: 111397, г. Москва, Зелёный проспект, д. 20  
Почтовый адрес: 111401, г. Москва, а/я 10

**Редакция журнала:**

Факт. адрес: 111397, г. Москва, Зелёный проспект, д. 20  
Почтовый адрес: 111401, г. Москва, а/я 10  
Тел.: +7 (495) 989-9610  
E-mail: post@fin-izdat.ru  
Website: http://www.fin-izdat.ru

Генеральный директор: **В.А. Горохова**Управляющий директор: **А.К. Смирнов**

Главный редактор: **В.А. Цветков**, доктор экономических наук, профессор,  
член-корреспондент РАН, Москва, Российская Федерация

Зам. главного редактора:

**В.В. Гаврилов**, Москва, Российская Федерация**Ю.В. Горбачева**, Москва, Российская Федерация**Редакционный совет:****А.М. Батьковский**, доктор экономических наук, профессор,  
Москва, Российская Федерация**А.З. Дадашев**, доктор экономических наук, профессор,  
Москва, Российская Федерация**В.Н. Едророва**, доктор экономических наук, профессор,  
Нижний Новгород, Российская Федерация**Г.Б. Клейнер**, доктор экономических наук, профессор, член-корреспондент РАН,  
Москва, Российская Федерация**О.П. Овчинникова**, доктор экономических наук, профессор,  
Белгород, Российская Федерация**С.В. Ратнер**, доктор экономических наук, доцент, Москва, Российская Федерация**А.Н. Сухарев**, доктор экономических наук, доцент, Тверь, Российская Федерация**Е.А. Федорова**, доктор экономических наук, профессор,  
Москва, Российская Федерация**И.В. Шевченко**, доктор экономических наук, профессор,  
Краснодар, Российская Федерация**Е.Ф. Сысоева**, доктор экономических наук, доцент,  
Воронеж, Российская ФедерацияОтветственный секретарь: **И.Л. Селина**Перевод и редактирование: **О.В. Яковлева, И.М. Вечканова**Веб-разработка: **А.А. Клюкин**Контент-менеджеры: **В.И. Романова, Е.И. Попова**Менеджмент качества: **А.Ю. Садкус, А.В. Бажанов**Верстка: **Н.И. Бранделис**Корректор: **Т.А. Кондратенко**Подписка и реализация: **Р.Р. Гуськова**

Подписано в печать 23.07.2015

Выход в свет 23.07.2015

Формат 60x90 1/8. Объем 8,5 п.л. Тираж 1 200 экз.

Отпечатано в ООО «КТК»

Юр. адрес: 141290, Российская Федерация, Московская обл., г. Красноармейск,

ул. Свердлова, д. 1

Тел.: +7 (496) 588-0866

**Подписка:**

Агентство «Урал-пресс»

Агентство «Роспечать» – индекс 71222

Объединенный каталог «Пресса России» – индекс 45029

Свободная цена

Журнал доступен в EBSCOhost™ databases

Электронная версия журнала: <http://elibrary.ru>, <http://dilib.ru>, <http://biblioclub.ru>

Точка зрения редакции не всегда совпадает с точкой зрения авторов публикуемых статей

Перепечатка материалов и использование их в любой форме, в том числе  
и в электронных СМИ, возможны только с письменного разрешения редакции  
Редакция приносит извинения за случайные грамматические ошибки

© ООО «Издательский дом ФИНАНСЫ и КРЕДИТ»

**СОДЕРЖАНИЕ****БАНКОВСКОЕ ДЕЛО**

- Шунина Ю.С., Алексеева В.А., Клячкин В.Н.*  
Прогнозирование кредитоспособности клиентов  
на основе методов машинного обучения 2
- Юдина И.Н.* Цена возможных ошибок  
регулирования капитала банков 13

**ДЕНЕЖНАЯ СИСТЕМА**

- Антонов П.В., Злобина О.О.* Оценка эффектив-  
ности денежно-кредитной системы Российской  
Федерации 25

**БЮДЖЕТНО-НАЛОГОВАЯ СИСТЕМА**

- Алиев Б.Х., Алимурзоева М.Г.* Макроэкономичес-  
кие факторы динамики налоговых доходов  
субъектов Российской Федерации 35
- Жарова Е.Н.* Особенности налогового стимулиро-  
вания отдельных групп инновационно активных  
предприятий в Российской Федерации 48

**СТОИМОСТЬ БИЗНЕСА**

- Муслиенко С.О.* Внедрение программно-целевого  
(нормативного) подхода в систему финансового  
планирования на малых предприятиях 59

# FINANCE & CREDIT

A peer reviewed analytical and practical journal

Since 1994  
4 issues per month

The journal is recommended by VAK (the Higher Attestation Commission) of the Ministry of Education and Science of the Russian Federation to publish scientific works encompassing the basic matters of theses for advanced academic degrees indexing in Referativny Zhurnal VINITI RAS. Included in the Russian Science Citation Index (RSCI) Registration Certificate ПИ № ФС77-60938 of March 02, 2015 by the Ministry of Press, Broadcasting and Mass Communications of the Russian Federation. Previous Registration Certificate № 013007 of November 03, 1994 by the Committee of the Russian Federation on Press

**Founder:**

Publishing house FINANCE and CREDIT  
Office: 111397, Zelenyi prospect 20, Moscow, Russian Federation  
Post address: 111401, P.O. Box 10, Moscow, Russian Federation  
Telephone: +7 495 989 9610

**Publisher:**

Information center Finance and Credit, Ltd.  
Office: 123182, Aviatsionnaya St. 79-480, Moscow, Russian Federation  
Post address: 111401, P.O. Box 10, Moscow, Russian Federation  
Telephone: +7 495 989 9610

**Editorial:**

Office: 111397, Zelenyi prospect 20, Moscow, Russian Federation  
Post address: 111401, P.O. Box 10, Moscow, Russian Federation  
Telephone: +7 495 989 9610  
E-mail: [post@fin-izdat.ru](mailto:post@fin-izdat.ru)  
Website: <http://www.fin-izdat.ru>

Director General: **Vera A. Gorokhova**  
Managing Director: **Aleksey K. Smirnov**

Editor-in-Chief: **Valerii A. Tsvetkov**, Market Economy Institute RAS,  
Moscow, Russian Federation

**Deputy Editors:**

**Vladimir V. Gavrilov**, Moscow, Russian Federation  
**Yuliya V. Gorbacheva**, Moscow, Russian Federation

**Editorial Council:**

**Aleksandr M. Bat'kovskii**, Central Research Institute of Economics, Control Systems and Information, Moscow, Russian Federation  
**Alikhan Z. Dadashev**, Academy of Public Administration, Moscow, Russian Federation  
**Valentina N. Edronova**, Lobachevsky State University of Nizhny Novgorod – National Research University, Nizhny Novgorod, Russian Federation  
**Georgii B. Kleiner**, Central Economics and Mathematics Institute, RAS, Moscow, Russian Federation  
**Oksana P. Ovchinnikova**, Belgorod National Research University, Belgorod, Russian Federation  
**Svetlana V. Ratner**, Trapeznikov Institute of Control Sciences RAS, Moscow, Russian Federation  
**Aleksandr N. Sukharev**, Tver State University, Tver, Russian Federation  
**Elena A. Fedorova**, Financial University under Government of RF, Moscow, Russian Federation  
**Igor' V. Shevchenko**, Kuban State University, Krasnodar, Russian Federation  
**Elena F. Syssoeva**, Voronezh State University, Voronezh, Russian Federation

**Executive Editor:** **Inna L. Selina**

Translation and Editing: **Olga V. Yakovleva**, **Irina M. Vechkanova**  
Web Development: **Anton A. Klyukin**

Content Managers: **Valentina I. Romanova**, **Elena I. Popova**  
Quality Management: **Aleksandr Yu. Sadkus**, **Andrey V. Bazhanov**  
Layout Designer: **Natal'ya I. Brandelis**  
Proofreader: **Tat'yana A. Kondratenko**  
Sales and Subscription: **Ravilya R. Gus'kova**  
Printed by KTK, Ltd., 141290, Sverdlov St., 1, Krasnoarmeysk, Russian Federation  
Telephone: +7 496 588 0866  
Published July 23, 2015. Circulation 1 200

**Subscription:**

Ural-Press Agency  
Rospechat Agency  
Press of Russia Union Catalogue

**Online version:**

EBSCOhost™ databases  
Scientific electronic library: <http://elibrary.ru>  
University Library Online: <http://biblioclub.ru>

Not responsible for the authors' personal views in the published articles

This publication may not be reproduced in any form without permission

All accidental grammar and/or spelling errors are our own

© Publishing house FINANCE and CREDIT

## CONTENTS

### BANKING

- Shunina Yu.S., Alekseeva V.A., Klyachkin V.N.*  
Forecasting the customers' creditworthiness through machine learning methods 2
- Yudina I.N.* Implications of possible errors in bank capital regulating 13

### MONETARY SYSTEM

- Antonov P.V., Zlobina O.O.* Evaluating the efficiency of the monetary system of the Russian Federation 25

### FISCAL SYSTEM

- Aliev B.Kh., Alimirzoeva M.G.* Macroeconomic factors of trends in tax revenue of regions of the Russian Federation 35
- Zharova E.N.* Specifics of tax incentives for some groups of innovation-driven companies in the Russian Federation 48

### BUSINESS VALUE

- Musienko S.O.* A performance-based (regulatory) approach to financial planning in small businesses 59

**ПРОГНОЗИРОВАНИЕ КРЕДИТОСПОСОБНОСТИ КЛИЕНТОВ  
НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ****Юлия Сергеевна ШУНИНА<sup>а</sup>, Венера Арифзяновна АЛЕКСЕЕВА<sup>б,\*</sup>, Владимир Николаевич КЛЯЧКИН<sup>с</sup>**

<sup>а</sup> аспирант кафедры прикладной математики и информатики, Ульяновский государственный технический университет, Ульяновск, Российская Федерация  
ydoncova@yandex.ru

<sup>б</sup> кандидат технических наук, доцент кафедры прикладной математики и информатики, Ульяновский государственный технический университет, Ульяновск, Российская Федерация  
v.a.alekseeva@bk.ru

<sup>с</sup> доктор технических наук, профессор кафедры прикладной математики и информатики, Ульяновский государственный технический университет, Ульяновск, Российская Федерация  
v\_kl@mail.ru

\* Ответственный автор

**История статьи:**

Принята 01.12.2014

Одобрена 17.02.2015

УДК 336.77

**Ключевые слова:**

кредитоспособность, машинное обучение, дискриминантный анализ, опорные векторы, логистическая регрессия

**Аннотация**

**Предмет и тема.** В статье рассматривается процесс прогнозирования кредитоспособности клиентов банка. В связи с ростом конкуренции на рынке кредитных услуг разработка новых элементов этого процесса и более точной оценки кредитного риска является актуальной задачей.

**Цели и задачи.** Целями работы являются совершенствование методики прогнозирования кредитоспособности клиентов на основе использования современных методов машинного обучения и формирование оптимального решения о выдаче кредита.

**Методология.** Предложен алгоритм прогнозирования кредитоспособности по известным характеристикам заемщика на основе методов машинного обучения (кластеризации, регрессионного анализа, классификации). Данный алгоритм позволяет использовать как отдельные модели, так и все возможные их комбинации. В рассматриваемом подходе также предлагается провести предварительный анализ данных (дискретизация, поиск статистически значимых характеристик заемщика) и использовать различные критерии качества для выбора оптимальной структуры. На основе полученных результатов клиенты банка по уровню кредитоспособности делятся на заданное число классов  $k$ .

**Результаты.** На основе разработанного алгоритма получен эффективный метод прогнозирования кредитоспособности, позволяющий оценивать вероятность выплаты кредита по известным характеристикам заемщика. Эффективность этого метода показана на примере. На основе 20 характеристик заемщика были построены различные модели классификации (как по отдельности, так и в виде различных их комбинаций), и среди них была выявлена структура с наименьшей среднеквадратической ошибкой прогнозирования. Разделение клиентов на более чем два класса позволяет оптимизировать процесс принятия решения по выдаче кредита в связи с уменьшением риска невозврата.

**Выводы и значимость.** Комбинация моделей с использованием машинного обучения улучшает результативность прогнозирования кредитоспособности, позволяет повысить качество оценки риска и оптимизировать процесс выдачи кредита.

© Издательский дом «ФИНАНСЫ и КРЕДИТ», 2015

**Постановка задачи**

Задачу прогнозирования кредитоспособности с точки зрения статистической теории принятия решений можно рассматривать как задачу классификации новых клиентов на основе информации о прошлых клиентах [1]. В области кредитования эта задача решается в рамках анкетного скоринга [2–5].

Формальная постановка данной задачи заключается в следующем. Пусть имеется множество клиентов банка  $\{Z_i\}$ ,  $i=1, \dots, n$ , каждый из которых характеризуется  $p$ -мерным вектором признаков  $X_i = (x_{i1}, \dots, x_{ip})^T$ . Известна также принадлежность каждого клиента  $Y_i$  к одному из двух классов кредитоспособности:

$$Y = \begin{cases} y = 1 - \text{клиент кредитоспособен;} \\ y = 0 - \text{клиент некредитоспособен.} \end{cases}$$

Соответствующая выборка является обучающей: на ее основе необходимо описать процедуры, с помощью которых можно было бы с наибольшей точностью отнести новых клиентов  $\{W_j\}$ ,  $j = 1, \dots, m$  к одному из классов  $k \geq 2$ , имея в качестве входной информации только наборы признаков  $X_j = (x_{j1}, \dots, x_{jp})^T$ , описывающих новых клиентов. Поскольку клиенты могут характеризоваться как количественными, так и качественными признаками, возникает задача классификации клиентов в пространстве разнотипных признаков.

Набор признаков  $x_1, \dots, x_p$ , как правило, представляет собой данные из анкет, которые заполняются при подаче заявки на кредит. Поскольку в положении Банка России<sup>1</sup> приводится лишь примерный перечень информации для анализа финансового положения заемщика, существуют различные формы анкет. Основной набор признаков: ФИО, дата рождения, паспортные данные, образование, адрес, семейное положение, ФИО и дата рождения родственников, сведения об основной работе, ежемесячные доходы и расходы, информация об имеющемся имуществе, информация об имеющихся долгах и обязательствах, информация об инвалидности, запрашиваемая сумма и т.д.

Помимо данных из анкеты, если клиент уже брал кредит в прошлом или имеет кредит в настоящее время, в качестве признаков могут также выступать данные о кредитной истории, полученные из кредитного бюро посредством запроса. Возможно несколько вариантов представления информации о кредитной истории. В первом случае это может быть обобщенное решение: кредитная история — положительная либо отрицательная. Отрицательной историей обычно считается история при задержке платежей сроком на три и более месяца. В другом случае в качестве признаков могут использоваться данные, полученные из кредитного бюро<sup>2</sup>.

Следует отметить, что один и тот же набор признаков может быть дан как кредитоспособным, так и некредитоспособным клиентом, поэтому принципиально невозможно достигнуть абсолютно точной классификации.

<sup>1</sup> Положение Банка России от 26.03.2004 № 254-П «О порядке формирования кредитными организациями резервов на возможные потери по ссудам, по ссудной и приравненной к ней задолженности».

<sup>2</sup> URL: <https://www.tcsbank.ru/tournament>.

Согласно положению Банка России от 26.03.2004 № 254-П «О порядке формирования кредитными организациями резервов на возможные потери по ссудам, по ссудной и приравненной к ней задолженности» финансовое положение заемщика может быть оценено как хорошее, среднее или плохое. При этом нет четких определений данной градации. Среди исследовательских работ в качестве выходной информации в основном используются два класса кредитоспособности. Однако столь жесткое разделение может привести к потере клиентов, которым можно было бы предложить другие условия кредитования, например меньший срок, более высокий процент и т.п. Поэтому в данной статье в качестве выходной информации предлагается использовать не только класс кредитоспособности, но и апостериорное распределение.

Апостериорное распределение указывает для каждого класса вероятность принадлежности клиента к этому классу. Например, в случае двух классов при распределениях 90%/10% и 55%/45% клиент будет кредитоспособен в обоих случаях, однако с очевидной разницей.

### Обзор методов и моделей для решения поставленной задачи

Для решения поставленной задачи применим ряд моделей классификации.

*Дискриминантный анализ* [6] — наиболее распространенный метод классификации при наличии обучающей выборки и заключается в использовании линейных скоринговых функций для определения вероятности принадлежности клиента к одному из  $k$  классов:

$$\begin{aligned} s_1(x) &= q_0^1 + q_1^1 x_1 + \dots + q_p^1 x_p; \\ s_2(x) &= q_0^2 + q_1^2 x_1 + \dots + q_p^2 x_p; \\ &\dots \\ s_k(x) &= q_0^k + q_1^k x_1 + \dots + q_p^k x_p, \end{aligned}$$

где  $x_1, \dots, x_p$  — набор признаков;

$q_0, \dots, q_p$  — параметры регрессии;

$s(x)$  — «счет», который содержит достаточное количество информации для того, чтобы различать класс клиента.

Выбирается тот класс, которому соответствует больший счет. В случае двух классов получается тот же результат, что и при линейной регрессии. При этом каждая переменная из класса должна быть

подчинена нормальному закону распределения.

*Логистическая регрессия* [7]. Делается предположение о том, что вероятность наступления события  $y = 1$  (клиент кредитоспособен) равна

$$\Pr\{y = 1 | X\} = f(z);$$

$$z = Q^T X = q_0 + q_1 x_1 + \dots + q_p x_p,$$

где  $X$  — вектор-столбец независимых переменных  $x_1, \dots, x_p$ ;

$Q$  — вектор-столбец параметров регрессии  $q_0, \dots, q_p$ ;

$f(z)$  — логистическая функция:

$$f(z) = \frac{1}{1 + e^{-z}}.$$

Поскольку  $Y$  принимает лишь значения 0 и 1, то вероятность второго возможного значения (клиент не кредитоспособен) равна

$$\Pr\{y = 0 | X\} = 1 - f(z) = 1 - f(Q^T X).$$

Таким образом, логистическая регрессия заменяет вероятность наступления события логарифмом шансов:

$$\begin{aligned} \log \frac{\Pr\{y = 1 | X\}}{\Pr\{y = 0 | X\}} &= \frac{f(z)}{1 - f(z)} = \\ &= q_0 + q_1 x_1 + \dots + q_p x_p = s(x). \end{aligned}$$

Для нахождения параметров  $q_0, \dots, q_p$  необходимо составить обучающую выборку, состоящую из множества пар  $\{(X_i, Y_i), i = 1, \dots, n\}$  (обучающих примеров). Обычно используется метод максимального правдоподобия, согласно которому выбираются параметры  $q$ , максимизирующие значение функции правдоподобия на обучающей выборке. Применяется метод градиентного спуска или метод Ньютона–Рафсона [8].

Для двух классов задача классификации решается следующим образом: новый клиент  $W_j$  является кредитоспособным ( $y = 1$ ), если предсказанная моделью вероятность  $\Pr\{y = 1 | X_j\} > 0,5$ , иначе клиент является не кредитоспособным ( $y = 0$ ). Граничное значение может быть отлично от 0,5. Логистическая регрессия является также адекватным математическим инструментом для оценки переходных вероятностей состояний марковских цепей<sup>3</sup>.

Недостатком как дискриминантного анализа, так и

<sup>3</sup> Клячкин В.Н., Донцова Ю.С. Сравнительный анализ точности нелинейных моделей при прогнозировании состояния системы на основе марковской цепи // Известия Самарского научного центра Российской академии наук. 2013. Т. 15. № 4. С. 924–927.

логистической регрессии является чувствительность к корреляции между факторами, поэтому в моделях недопустимо наличие сильнокоррелированных зависимых переменных. Кроме того, оба подхода не могут использоваться для нелинейных процессов. К их преимуществам можно отнести возможность проведения исследования и взвешивания факторов, влияющих на результат. Более того, логистическая регрессия учитывает ограничения на значения вероятности, которые не могут выходить за рамки 0 и 1.

*Метод нейронных сетей.* Нейронные сети представляют собой систему соединенных и взаимодействующих между собой искусственных нейронов [9]. Каждый нейрон состоит из трех основных компонентов — синапсов (умножителей), сумматора, нелинейного преобразователя. С помощью синапсов происходят передача сигнала и умножение его на весовой коэффициент, который определяет силу связи. Сумматор складывает сигналы, поступающие от синапсов, а последний компонент производит преобразование просуммированного сигнала согласно некоторой функции активации нейрона.

Математическая модель нейрона имеет следующий вид:

$$S = \sum_{r=1}^p q_r x_r + b, \quad Y = F(S),$$

где  $S$  — результат суммирования;

$q_r$  — вес  $r$ -го синапса;

$x_r$  — входной сигнал;

$b$  — значение смещения;

$Y$  — выходной сигнал;

$p$  — число входов нейрона;

$F$  — функция активации.

С математической точки зрения обучение нейронных сетей — это многопараметрическая задача нелинейной оптимизации, которая заключается в нахождении коэффициентов связей между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение. Это значит, что в случае успешного обучения сеть сможет вернуть верный результат на основании как отсутствующих, так и неполных, а также частично искаженных данных.

Нейронные сети чаще применяются для скоринга юридических лиц, чем для физических. Также

нейронные сети наилучшим образом проявили себя при выявлении мошенничества с кредитными карточками, поскольку они способны указывать на нестандартные ситуации [10].

К преимуществам нейронной сети, помимо возможности ее применения на неполных исходных данных, относятся способность устанавливать нелинейные связи между прогнозными и фактическими значениями процессов, а также возможность быстрой адаптации к изменяющимся внешним условиям. Главным недостатком сети является то, что веса связей, определенные в результате обучения, не имеют никакой интерпретации в терминах кредитного риска. Следовательно, практически невозможно объяснить предсказание, а также провести анализ чувствительности, чтобы выделить наиболее значимые параметры. При этом выбор количества входных факторов существенно влияет на время обучения.

*Деревья принятия решений* [11] последовательно разделяют клиентов на классы по одной из переменных так, чтобы эти классы максимально возможно отличались по величине кредитного риска. При этом на первом шаге деление производится по самому значимому фактору. Далее процесс продолжается до того момента, пока оставшиеся классы не становятся настолько малы, что следующее разбиение не приведет к статистически значимому различию на уровне риска. Количество классов на каждом шаге процедуры построения дерева решений выбирается автоматически.

К преимуществам деревьев решений относятся быстрая обработка больших объемов данных, легкая интерпретируемость результатов, работа с пропущенными, числовыми и нечисловыми типами данных, а также отсутствие ограничений на коррелируемость между зависимыми переменными. К недостаткам метода можно отнести неоднозначность алгоритма построения структуры дерева, а также вопрос о том, когда стоит прекратить дальнейшее деление на классы.

*Метод опорных векторов* [12]. Основная идея метода заключается в переводе исходных векторов в пространство более высокой размерности, а также в поиске оптимальной гиперплоскости, разделяющей классы клиентов наилучшим образом. На обучающей выборке  $\{(X_i, Y_i), i = 1, \dots, n\}$  метод опорных векторов строит классифицирующую функцию, которая имеет следующий вид:

$$F(X) = \text{sign}(\langle v, X \rangle + b),$$

где  $\langle \cdot, \cdot \rangle$  — скалярное произведение;

$v$  — нормальный вектор к разделяющей гиперплоскости;

$X$  — вектор-столбец признаков;

$b$  — вспомогательный параметр.

Поскольку расстояние от оптимальной гиперплоскости до класса должно быть максимально, для нахождения  $v$  и  $b$  возникает задача оптимизации:

$$\begin{cases} \arg \min_{v, b} \|v\|^2; \\ Y_i(\langle v, X_i \rangle + b) \geq 1, i = 1, \dots, n. \end{cases}$$

Данная задача решается, как правило, методом множителей Лагранжа [13].

Результат классификации новых клиентов находится следующим образом: при  $F(X) = 1$  клиент считается кредитоспособным, при  $F(X) = 0$  — некредитоспособным.

Преимущество метода заключается в том, что оптимизация здесь является задачей квадратичного программирования в выпуклой области, которая всегда имеет единственное решение. При этом ищется разделяющая полоса максимальной ширины, что позволяет в дальнейшем уверенно классифицировать новых клиентов. Недостатком можно назвать то, что метод чувствителен к шумам и стандартизации данных, а также отсутствует общий подход к автоматическому выбору ядра в случае линейной неразделимости классов.

*Наивный байесовский классификатор* [14] представляет собой граф, который можно условно разделить на две части. Первая часть представляет собой вершину класса кредитоспособности клиента  $K$ , а вторая часть — вершины независимых переменных  $x_1, \dots, x_p$ , к которым проведены стрелки из вершины класса.

Обучение классификатора означает оценку условных вероятностей  $P(X | K)$ , а классификация на  $k$  классов производится простым применением формулы Байеса:

$$P(K = k | X = x) = \frac{P(K = k) \prod_{r=1}^p P(X_r = x_r | K = k)}{\sum_{k'} P(K = k') \prod_{r=1}^p P(X_r = x_r | K = k')}$$