

А.Г. Шмелева¹, И.В. Каленюк¹, С.Ю. Обыденнова²,
Д.В. Сяиков³, Г.Д. Пильщиков³, С.С. Шилков¹

ПРОГРАММНАЯ МОДЕЛЬ ОЦЕНКИ КРЕДИТОСПОСОБНОСТИ КЛИЕНТОВ С ПРИМЕНЕНИЕМ АЛГОРИТМОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

МИРЭА – Российский технологический университет¹
Нижегородский государственный технический университет им. Р.Е. Алексеева²
Российский университет дружбы народов³

Обоснована актуальность применения методов искусственного интеллекта для оценки кредитоспособности физических лиц. Построена математическая модель оценки кредитоспособности посредством присвоения физическому лицу соответствующего коэффициента доверия, при этом отражены лишь некоторые оцениваемые характеристики заемщика для формирования скоринг-балла. На основании построенной математической модели реализована программа на языке программирования С#. Представлены результаты численного моделирования оценки кредитоспособности клиента с применением разработанной программной реализации на основе предложенной скоринговой модели. Произведена оценка кредитоспособности физических лиц, полученные результаты были сравнены с оценкой «Национального бюро кредитных историй». Разработанная модель и программа может применяться для оценивания и ранжирования объектов в разнообразных прикладных задачах имитационного моделирования. Программа имеет простой интерфейс, возможность выбора параметров оценки, расширения их числа. Реализованные алгоритмы машинного обучения позволяют уточнять коэффициенты модели с ростом статистических данных.

Ключевые слова: имитационное моделирование, машинное обучение, скоринг-балл, кредитный скоринг, оценка кредитоспособности.

Введение

В условиях цифровой экономики актуальным является применение алгоритмов машинного обучения и искусственного интеллекта для анализа данных о клиентах с целью получения знаний, снижающих риски организации при заключении договоров. Наличие информации о клиенте, его предпочтения, потребительские запросы и финансовые возможности, позволяют лицу, принимающему решения формировать специальные предложения, ориентированные на определенный тип клиентов, разрабатывать стратегии увеличения числа клиентов с минимизацией рисков, способствующих эффективному развитию организации. Разрабатываемые методы оценки клиентов способствуют составлению объективных оценок, что обеспечивает снижение рисков при принятии решений. Например, в настоящее время статистические методы широко применяются банками для оценки кредитоспособности заемщиков. Данные методы объединены термином «кредитный скоринг» – каждому клиенту на основе анализа персональных данных ставится в соответствие скоринг-балл, который упрощает и автоматизирует процесс первоначального принятия решений по предоставлению или отказу в кредите.

Информационные системы, включающие инструменты скоринга, имеют ряд преимуществ: скорость, систематизация и объективность оценки характеристик клиентов. Например, возрастающая потребность населения в кредитных средствах, а следовательно, и возросшая из-за этого конкуренция между банковскими учреждениями, требует от банков создание и развитие собственных уникальных скоринговых моделей для оптимизации принятия решений по заявкам с учетом критериев валидности клиента, которые в разных организациях могут отличаться. Первоначальные варианты скоринговых моделей базировались на линейных методах, учитывающих весовые коэффициенты, но развитие информационных

технологий требует совершенствования внутрибанковских скоринг-систем с учетом собственных накапливаемых данных о клиентах.

В современном мире каждый час происходит огромное количество операций прямо или косвенно связанных со скорингом; например, скоринговые системы применяются не только в оценке кредитоспособности заемщика, а также в оценке причин уменьшения количества клиентов, в области увеличения кредитного лимита по кредитным картам или оценки просроченных платежей, анализируя причины пропуска. Широкий спектр услуг и большой объем операций требует создание инструментов, направленных на анализ данных, способствующих принятию оптимального решения. Например, еще недавно для оформления кредита обязательно было личное присутствие в банке заемщика для предоставления документов, что, в свою очередь, накладывало определенную нагрузку на штат сотрудников, время на обработку данных вручную и, в итоге, порождало потерю времени, средств и клиентов. Процедура рассмотрения и одобрения (отказа) кредита занимала долгое время – это в современных реалиях неэффективно, так как приводит к потере клиентов, а также со стороны банковского сектора к дополнительным расходам на содержание сотрудников, занимающихся обработкой данных вручную. Цифровая экономика направлена на внедрение информационных систем [1-3], способствующих автоматизации принятия решений, повышающих точность и быстродействие управления. Например, развитие интернет банкинга позволяет упростить способы взаимодействия с банками, в частности, автоматический анализ кредитоспособности и подбор наилучшего варианта, как для банка, так и для клиента с целью уменьшения рисков обеих сторон.

В течение последних лет в банковском секторе наблюдается увеличение объема размещенных денежных средств. Объединенное Кредитное Бюро (ОКБ) подвело итоги кредитной активности граждан за I квартал 2019 г. на основе информации от 600 кредиторов, передающих данные в бюро [4]. Всего за I квартал 2019 г. банки выдали россиянам 8,84 млн новых кредитов общим объемом свыше 1,97 трлн руб. В годовом отношении количество выданных кредитов выросло на 11 %, при этом объемы кредитования увеличились на 17 %. В I квартале 2018 г. было выдано 7,95 млн кредитов на 1,67 трлн руб. Количество новых кредитов наличными выросло на 16 %, при этом объем кредитования увеличился на 31 %. За январь-март 2019 г. было выдано 3,71 млн кредитов наличными на 998,35 млрд руб. против 3,20 млн кредитов на 763,45 млрд руб. годом ранее. Средний размер выданного кредита наличными вырос на 12 %: с 239 до 269 тыс. руб. Кредитная нагрузка россиян повышается с каждым годом, что ставит перед банками задачи разработки инструментов принятия решений по оценке кредитоспособности клиентов с учетом их возможностей для минимизации рисков как банка, так и клиента. Это обосновывает необходимость разработки методов и алгоритмов [5], которые позволят быстро и эффективно на основании предоставленных данных со стороны заемщика, с применением внедряемых информационных систем, включающих инструменты мониторинга и статистики, принимать решение о выдаче кредита заемщику с целью минимизации взаимных рисков.

Математическая модель кредитного скоринга

Современные скоринг-модели могут включать набор методов и алгоритмов анализа данных: регрессионный анализ, нейронные сети, методы машинного обучения и искусственного интеллекта, обладающих следующими преимуществами: анализ большого количества признаков, установление закономерностей, быстродействие. Стоит учитывать, что для эффективной классификации и принятия решений необходимо осуществлять накопление статистических данных о клиентах, которые позволят произвести формирование обучающей выборки и в дальнейшем уточнение весовых коэффициентов.

Полученные при анализе данных знания о клиентах формируют стратегию развития компании посредством разработки предложений направленных на определенный тип клиентов в результате анализа и классификации статистических данных. Для демонстрации оценки

кредитоспособности клиентов рассмотрим построенную скоринг-модель анализа кредитной заявки клиента при программной реализации которой применялись нейронные сети. Созданная программа позволяет при накоплении данных вносить дополнительные критерии оценки.

Цель модели – классификация данных об объекте исследования, в рассматриваемой задаче о клиенте (физическом лице) согласно его характеристикам и соответствующим весам, определенным на основе анализа статистических данных. При накоплении статистических данных значения весовых коэффициентов уточняются в процессе эксплуатации созданной программы.

В общем случае для обученной нейронной сети можно установить бинарное значение выходных переменных. Например, для разработанной и тестируемой скоринг-модели достаточно двух классов результата:

- клиенту можно одобрить заявку и заключить договор;
- клиенту нельзя одобрить заявку и, следовательно, нельзя заключить договор.

По принципу построения нейронной сети, зададим функцию активации нейрона – пороговую функцию. Поскольку существует всего два результирующих класса, например, можно кодировать их при помощи 0 и 1 (где 0 – нельзя заключить договор, 1 – можно заключить договор).

Для реализации данной модели мы будем использовать пороговую функцию (1):

$$f(x) = \begin{cases} 1, & \text{если } x \geq d, \\ 0, & \text{если } x < d, \end{cases} \quad (1)$$

где d – это коэффициент доверия, пороговое значение, для нашей модели примем $d = 0,7$. Таким образом, будем считать, что объект оценки на 70 % удовлетворяет установленным критериям. Пороговое значение d вычисляется в зависимости от оценки характеристик клиента. Например, наличие действующих кредитов, заработная плата, образование и т.д. Для оценки характеристик клиента формируется список оцениваемых параметров и соответствующие им начальные значения весовых коэффициентов, в процессе обучения коэффициенты уточняются. На основе статистических данных числовые значения характеристик нормированы в диапазоне значений от 0 до 1.

В рассматриваемом модельном примере приведены оценки базовых характеристик клиента для принятия эффективных кредитных решений. Предложенная модель позволяет использовать как бинарные, так и небинарные переменные. Например, бинарные переменные: есть ли за последний год пропуски платежей по кредиту 0 – нет, 1 – да. По результатам анализа статистических данных о клиентах определены начальные значения весовых коэффициентов. При изменении статистической выборки данные коэффициенты определяются в процессе обучения. При построении первоначальной модели учитывалось, что есть значимые характеристики, например, лучше работать с платежеспособными клиентами, поэтому для данной характеристики присвоено начальное значение весового коэффициента равное 0,7. По аналогии формировались и другие значения.

Для демонстрации работы созданной программы представим модельный пример, содержащий список некоторых главных оцениваемых характеристик клиента.

1. Наличие двух и более просроченных платежей по кредитам за год – бинарный параметр, принимающий значение 0 или 1. Вес $w_1 = -0,8$.

2. Отсутствие просроченных платежей по кредитам за год – бинарный параметр, принимающий значение 0 или 1. Вес $w_1 = 0,7$.

3. Заработная плата клиента, где за y принимаем минимальный размер оплаты труда в регионе проживания клиента (2):

$$f(x) = \begin{cases} 1, & x \geq 4 * y; \\ 0,7, & 2 * y < x \leq 3 * y; \\ 0,5, & y < x \leq 2 * y; \\ 0,3, & 0,5 * y < x \leq y \\ 0, & x \leq 0,5 * y. \end{cases} \quad (2)$$

Присвоим вес $w_3 = 1$.

4. Кредитная нагрузка клиента, в данном случае будем учитывать соотношение ежемесячных платежей по кредитам к доходу клиента (3):

$$f(x) = \begin{cases} 1, & 0,7 < x; \\ 0,7, & 0,5 < x \leq 0,7; \\ 0,5, & 0,3 < x \leq 0,5; \\ 0,3, & 0,1 < x \leq 0,3; \\ 0, & x \leq 0,1. \end{cases} \quad (3)$$

Вес $w_4 = -1$. Значение весового коэффициента присвоено исходя из предположения, что данный фактор, как и заработанная плата клиента, играет важную роль общей модели.

5. Стаж работы клиента в организации (4):

$$f(x) = \begin{cases} 1, & x \text{ более } 10 \text{ лет;} \\ 0,7, & x \text{ в диапазоне } 7 - 10 \text{ лет;} \\ 0,5, & x \text{ в диапазоне } 4 - 6 \text{ лет;} \\ 0,3, & x \text{ в диапазоне } 1 - 3 \text{ лет;} \\ 0, & x \text{ менее } 1 \text{ года.} \end{cases} \quad (4)$$

Вес $w_5 = 0,5$.

6. Возраст заемщика (5):

$$f(x) = \begin{cases} 1, & x \text{ в диапазоне } 33 - 40; \\ 0,7, & x \text{ в диапазоне } 27 - 32 \text{ и } 41 - 47; \\ 0,5, & x \text{ в диапазоне } 22 - 26 \text{ и } 48 - 55; \\ 0,2, & x \text{ в диапазоне } 18 - 22 \text{ и } 56 - 59; \\ 0,1, & x \text{ от } 60 \text{ и старше.} \end{cases} \quad (5)$$

Вес $w_6 = 0,7$.

7. Образование заемщика (6):

$$f(x) = \begin{cases} 1, & \text{высшее;} \\ 0,5, & \text{среднее специальное;} \\ 0,3, & \text{среднее;} \\ 0, & \text{остальное.} \end{cases} \quad (6)$$

Вес $w_7 = 0,7$.

8. Условия труда, будем классифицировать по степени вредности (7):

$$f(x) = \begin{cases} 1, & \text{оптимальные условия труда;} \\ 0,7, & \text{допустимые условия труда;} \\ 0,3, & \text{вредные условия труда;} \\ 0, & \text{опасные условия труда.} \end{cases} \quad (7)$$

Вес $w_8 = 0,5$.

9. Наличие задолженности по данным Федеральной службы судебных приставов (ФССП) – бинарный параметр, принимающий значения 0 или 1.

Вес $w_9 = -0,6$.

10. Наличие аккаунта в социальных сетях – бинарный параметр, принимающий значения 0 или 1.

Вес $w_{10} = 0,5$.

11. Наличие стационарного телефона – бинарный параметр, принимающий значения 0 или 1.

Вес $w_{11} = 0,4$.

Рассмотренный модельный пример содержит некоторые главные характеристики клиента, в созданной программе на усмотрение лица принимающего решения можно добавлять дополнительные пункты. Использование созданной программы позволяет провести быструю и эффективную оценку клиента по вопросу одобрения или отказа кредитной заявки.

Программная реализация и результаты расчетов

В качестве языка программирования был выбран язык C# в среде разработки Visual Studio. При открытии программы автоматически загружаются весовые коэффициенты, которые могут быть, как заранее подготовлены, так и динамически обновляться на основе данных о новых клиентах. Окно программы (рис. 1) позволяет ввести необходимую информацию о клиенте и на ее основе рассчитать скоринг-балл. Помимо этого, с помощью пункта меню «Действия → Добавить дополнительную информацию» в программе предусмотрена возможность добавления необязательной информации о клиенте, которую можно использовать для уточнения весовых коэффициентов и усовершенствования параметров скоринговой модели. Программа содержит алгоритмы машинного обучения на основе нейронных сетей, что позволяет при накоплении статистических данных формировать обучающую и тестовую выборку, расширять оцениваемые параметры и повышать точность классификации оценки кредитоспособности клиентов.

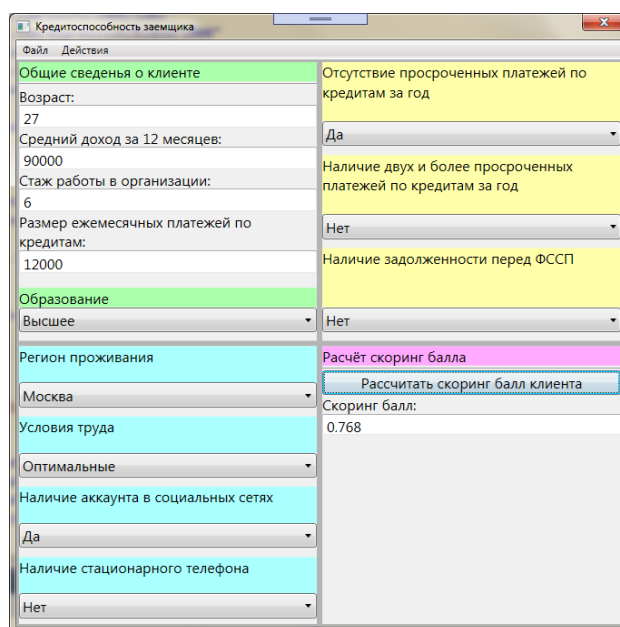


Рис. 1. Основное окно программы

Рассмотрим пример оценки кредитоспособности заемщиков по предлагаемой скоринг-модели с помощью разработанной программы. В табл. 1 представлена информация о четырех клиентах. Предполагается, что все клиенты проживают в Москве.

Таблица 1.

Данные о клиентах

	Клиент 1	Клиент 2	Клиент 3	Клиент 4
Просроченные платежи по кредитам за год	Нет	Нет	Да	Нет
Отсутствие пропуска платежей за год	Да	Да	Нет	Да
Доход, руб.	90 000	19 000	45 000	21 000

Окончание табл. 1.

	Клиент 1	Клиент 2	Клиент 3	Клиент 4
Кредитная нагрузка	Платежи по кредитам 12000, т.е., нагрузка слабая = 0,3	Нет	Платежи по кредитам 4500, т.е. нагрузка слабая = 0,3	Платежи по кредитам 8000, т.е., нагрузка = 0,5
Стаж работы в организации	6 лет	1 год	17 лет	Не работает
Возраст заемщика	27 лет	19 лет	45 лет	67 лет
Образование заемщика	Высшее	Неоконченное высшее	Высшее	Среднее специальное
Условия труда	Отличные	Отличные	0,7	Не работает
Наличие задолженности ФССП	Нет	Нет	Есть	Нет
Наличие аккаунта в социальных сетях	Есть	Есть	Есть	Нет
Наличие стационарного телефона	Нет	Нет	Нет	Есть

В табл. 2 представлены результаты оценки параметров клиентов. В последней строчке для каждого из них записан скоринг-балл, определенный с помощью разработанной программы.

Таблица 2.

Результаты анализа

Параметр	Вес параметра	Клиент 1	Клиент 2	Клиент 3	Клиент 4
Просроченные платежи	-0,8	0	0	1	0
Отсутствие пропуска платежей	0,7	1	1	0	1
Заработная плата	1	1	0,3	0,7	0,3
Кредитная нагрузка	-1	0,3	0	0,3	0,5
Стаж работы в организации	0,5	0,5	0,3	1	0
Возраст заемщика	0,7	0,7	0,2	1	0,1
Образование заемщика	0,7	1	0,3	1	0,5
Условия труда	0,5	1	1	0,7	0
Наличие задолженности ФССП	-1	0	0	1	0
Наличие аккаунта в социальных сетях	0,5	1	1	1	0
Наличие стационарного телефона	0,4	0	0	0	1
Скоринг-балл		0,768	0,5	0,27	0,264

Рассмотрев имеющиеся параметры, можно сделать следующие выводы.

Клиент 1 является усредненным примером положительного клиента для кредитора, наличие высокого дохода, отсутствие пропуска платежей и низкая кредитная нагрузка положительно сказалась на общей оценке.

Клиент 2 не имеет кредитной нагрузки, а также пропуска платежей, но финансовое положение не стабильное, маленький стаж работы, это отобразено в итоговой оценке.

Клиент 3 получает большой доход, а также относится к выгодной возрастной категории в предложенной системе оценке, но получил схожую оценку с клиентом 4, который имеет соответственно меньший доход и относится к не столь выгодной возрастной категории.


Связано это с наличием у клиента 3 пропуска платежей и задолженности по данным ФССП, что негативно сказалось на общей оценке.

Стоит отметить, что необходимо накапливать статистические данные для формирования обучающей выборки – это позволит произвести уточнение весовых коэффициентов. Для более эффективно анализа клиента аналитику необходимо дополнять количество рассматриваемых характеристик. Однако, даже рассмотрев не столь большое количество имеющихся параметров можно сделать определенную оценку.

Результаты тестирования работы программы совпали с данными НБКИ [3]. Пример для одного клиента представлен в табл. 3.

Таблица 3.

Сравнение результатов

Скоринг-балл, полученный в результате расчета программы	Скоринг-балл на основании отчета НБКИ
<div style="border: 1px solid gray; padding: 5px; width: fit-content;"> <p>Скоринг балл:</p> <p>0,795</p> </div>	

Сравнение результатов с учетом основных характеристик показало соответствие, отличие обусловлено различием в принимаемом максимальном значении скоринг-балла – максимальное значение кредитного рейтинга в отчете НБКИ 850 баллов, а в разработанной модели и программе 1000. Предложенная скоринг-модель и созданная программа обладают рядом преимуществ. Например, при деятельности банковской организации веса уточняются при помощи алгоритмов машинного обучения. Если был заключен договор с клиентом (в соответствии со скоринг-моделью), но при этом сотрудничество не принесло ожидаемых результатов, веса в модели меняются при обучении.

Заключение

В условиях цифровой экономики актуальным является разработка и применение инструментов автоматизации бизнес-процессов и принятия решений, например, оценки клиентов. Наличие информации о клиенте позволяет оценить его предпочтения, потребительские запросы и возможности. Применение искусственного интеллекта для оценки кредитоспособности заемщиков позволяет анализировать множество дополнительных критериев, выявляя индивидуальные предпочтения каждого клиента, финансовое положение, круг общения, потребительские запросы. Все это позволяет построить полную картину о клиенте и учесть все риски для кредитора.

В свою очередь, для банка это несет следующие преимущества.

1. Снижение штата сотрудников отдела проверки и рассмотрения кредитных заявок, а, следовательно, и расходов на данные отделы.
2. Уменьшение времени на рассмотрение заявки на кредитный продукт.
3. При анализе и принятии решении по заявке клиента не принимает прямого решения сотрудник, это позволяет избежать влияния человеческого фактора.
4. В зависимости от политики компании допустимо изменять веса значимости характеристик клиента, производить имитационное моделирование возможных альтернатив, на основе которых принимать стратегические решения управления.

Предложенная математическая модель и прикладная программа позволяют автоматизировать процессы анализа и оценки кредитных заявок, поступающих в банки, что повышает точность и быстрдействие принимаемых сотрудниками банковской организации решений.

Также разработанная модель и программа может применяться для оценивания и ранжирования объектов в разнообразных прикладных задачах имитационного моделирования. Программа имеет простой и ясный интерфейс, возможность выбора параметров оценки, расширения их числа. Реализованные алгоритмы машинного обучения позволяют уточнять коэффициенты модели с ростом статистических данных.

Библиографический список

1. **Shmeleva, A.G.** Industrial Management Decision Support System: from Design to Software. Proceedings of the 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus) / A.G. Shmeleva, A.I. Ladynin. – 2019. – P. 1474-1477. DOI: 10.1109/EIConRus.2019.8657313.
2. **Тюлин, А.Е.** Интеллектуальная космическая система в цифровой экономике: новые возможности и перспективы / А.Е. Тюлин, А.А. Чурсин, Р.В. Шамин, А.В. Юдин // Современная научная мысль. – 2018. – № 6. – С. 213-223.
3. **Shmeleva, A.G.** Manufacturing Planning Information System Development / A.G. Shmeleva, Y.V. Talanova, E.A. Galemina, A.I. Ladynin // Proceedings of the 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). – P. 366-369. DOI: 10.1109/EIConRus.2018.8317108.
4. Объединенное кредитное бюро. <https://bki-okb.ru/press/news/v-i-kv-2019-g-zaemshchiki-vzyali-kreditov-na-197-trln-rub>.
5. **Шмелева, А.Г.** Когнитивное моделирование в информационной системе поддержки принятия решений / А.Г. Шмелева, Ю.В. Таланова, В.В. Наумов, А.И. Ладынин // Труды НГТУ им. Р.Е. Алексеева. Нижний Новгород. – 2018. – Т. 78. – № 2 (121). – С. 60-67.

*Дата поступления
в редакцию: 03.08.2020*

**A.G. Shmeleva¹, I.V. Kalenyuk¹, S.Yu. Obydenнова²,
D.V. Syasikov³, G.D. Pilshchikov³, S.S. Shilkov¹**

A SOFTWARE MODEL FOR ASSESSING CUSTOMERS' CREDITWORTHINESS USING ARTIFICIAL INTELLIGENCE ALGORITHMS

MIREA – Russian technological university¹
Nizhny Novgorod State Technical University n.a. R.E. Alekseev²
Russian University of Peoples³

Purpose: Current trends in the development of digital infrastructures organizations in digital economy conditions are directed on the development of decision-making tools based on data analysis.

Design/methodology/approach: Organizations designed information systems must contain relevant elements that include automation of decision-making without human participation. The article substantiates the relevance of use of artificial intelligence methods for evaluating the creditworthiness of individual person.

Findings: A mathematical model for assessing creditworthiness was constructed by assigning a corresponding confidence coefficient to an individual. The program is implemented in the C# programming language.

Research limitations/implications: The results of numerical modeling of the client's creditworthiness assessment using the developed software implementation based on the proposed scoring model are presented. An assessment of the creditworthiness of individuals was made, the results were compared with the assessment of the «National Bureau of Credit Histories».

Originality/value: The developed model and program can be used for evaluating and ranking objects in various simulation applications. The program has a clear interface, the ability to select evaluation parameters and expand their number. Implemented machine learning algorithms allow you to refine the coefficients of the model with the growth of statistical data.

Key words: simulation modeling, machine learning, scoring score, credit scoring, credit rating.