
**МАТЕМАТИЧНІ МЕТОДИ, МОДЕЛІ
ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В ЕКОНОМІЦІ**

УДК 519.86

Біленко Д.В.кандидат економічних наук,
доцент кафедри бізнес-статистики та економічної кібернетики
Донецького національного університету імені Василя Стуса**Сивицька І.Г.**кандидат економічних наук,
доцент кафедри бізнес-статистики та економічної кібернетики
Донецького національного університету імені Василя Стуса**Теленкова Д.Г.**магістрант
Донецького національного університету імені Василя Стуса**МОДЕЛЮВАННЯ ОЦІНКИ КРЕДИТОСПРОМОЖНОСТІ
КЛІЄНТІВ БАНКУ ЗА ДОПОМОГОЮ МЕТОДУ ШТУЧНОГО ІНТЕЛЕКТУ**

У статті розглянуто наявні методи оцінки кредитоспроможності позичальника – фізичної особи, визначено їх недоліки та переваги. На основі аналізу недоліків наявних методів сформульовано вимоги до моделі оцінки кредитоспроможності клієнтів банку. Запропоновано методику моделювання оцінки кредитоспроможності клієнтів банку за допомогою методу штучного інтелекту, що полягає у виокремленні та аналізі великої кількості неструктурованої інформації з використанням статистичних та математичних методів для отримання нових, структурованих даних щодо кредитоспроможності клієнтів банку. Модель реалізовано на реальних даних та визначено її переваги, які полягають у можливості навчити наявну модель визначати нижню межу кредитоспроможності на основі реальних характеристик клієнтів, що істотно підвищує ефективність визначення неплатоспроможних клієнтів банку.

Ключові слова: кредитоспроможність клієнтів банку, скорингова оцінка, метод штучного інтелекту.

В статье рассмотрены существующие методы оценки кредитоспособности заемщика – физического лица, определены их недостатки и преимущества. На основе анализа недостатков существующих методов сформулированы требования к модели оценки кредитоспособности клиентов банка. Предложена методика моделирования оценки кредитоспособности клиентов банка с помощью метода искусственного интеллекта, который заключается в выделении и анализе большого количества неструктурированной информации с использованием статистических и математических методов для получения новых, структурированных данных о кредитоспособности клиентов банка. Модель реализована на реальных данных, определены ее преимущества, которые заключаются в возможности научить существующую модель определять нижнюю границу кредитоспособности на основе реально существующую характеристик клиентов, что существенно повышает эффективность определения неплатежеспособных клиентов.

Ключевые слова: кредитоспособность клиентов банка, скоринговая оценка, метод искусственного интеллекта.

Постановка проблеми. У сучасних умовах функціонування економіки однією з основних умов забезпечення фінансової безпеки банку є наявність ефективної системи аналізу та оцінки кредитоспроможності позичальників. Відповідно до статистики НБУ, банківська система України в 2017 р. встановила світовий рекорд за розміром проблемних кредитів, що становить майже 1 трлн. грн., а в березні 2018 р. цей показник сягнув 179 млрд. грн. [5]. Осно-

вними чинниками виникнення проблемної заборгованості є неповна або недостовірна інформація про позичальника, недостатньо продумана і розроблена кредитна політика банку, неякісна оцінка кредитоспроможності позичальника, недостатній рівень контролю позичальника після видачі кредиту, великий сегмент пов'язаних кредитів, які видавалися завідомо без мети їх повернення. Вищезазначене зумовлює актуальність наукових та

практичних досліджень у сфері розроблення моделей оцінки кредитоспроможності позичальників банку.

Аналіз останніх досліджень і публікацій. Дослідженню різних аспектів оцінки кредитоспроможності клієнтів банку присвячено роботи Я.В. Вовка, У.В. Хмеленко, Н.В. Бабіної, І.В. Ворошилової, Т.Б. Стечишина та ін. Проблеми моделювання кредитоспроможності позичальників комерційного банку розглядалися у працях В.В. Вітлінського, Я.С. Наконечного, Г.М. Крістіогло, Г.І. Великоіваненко. Разом із тим питання використання інтелектуального аналізу даних для визначення кредитоспроможності клієнтів банку потребує подальшого дослідження.

Метою дослідження є моделювання оцінки кредитоспроможності клієнтів банку за допомогою методу штучного інтелекту.

Виклад основного матеріалу дослідження. Діяльність будь-якої організації передбачає наявність різних видів ризиків, для комерційних банків одним з основних є кредитний ризик. Кредитний ризик – це вірогідність невиконання боржником своїх банківських зобов'язань. Нині для зниження кредитного ризику проводиться аналіз інформації про майбутніх клієнтів-позичальників різноманітними методами та моделями.

Оцінка кредитоспроможності позичальника – фізичної особи може проводитися на основі кількісних показників і якісних характеристик, серед яких на практиці найчастіше використовуються: скорингова оцінка, оцінка на основі фінансових показників платоспроможності, кредитна історія, андеррайтинг [9].

Одним із найбільш популярних методів оцінки кредитоспроможності клієнтів банку є скоринговий метод – діагностика ймовірності банкрутства потенційного позичальника під час розгляду питання можливості його кредитування. Скоринг передбачає класифікацію або розподіл бази клієнтів на різноманітні групи. В основі скорингових систем лежить припущення, що люди зі схожими соціальними показниками поведуться однаково.

Оцінка на основі фінансових показників здійснюється за допомогою розрахунків коефіцієнту платежу до доходу, коефіцієнту зобов'язання до доходу, інтегрального коефіцієнту платоспроможності позичальника – фізичної особи [3]. Отримані розрахункові значення порівнюються з нормативними, на основі чого робиться висновок щодо платоспроможності клієнта банку.

Методика визначення платоспроможності є дещо схожою з оцінкою на основі фінансових показників, де платоспроможність визна-

чається на підставі документів із місця роботи про доходи та розміри утримань, а також за даними анкети. Результат обчислюється як середньомісячний дохід за вирахуванням усіх обов'язкових платежів, скоригований на поправочний коефіцієнт і помножений на термін кредиту. Виходячи з отриманої суми, розраховується максимальний розмір кредиту [4].

Андеррайтинг є методом оцінки можливостей клієнта з обслуговування кредиту або розрахунок суми кредиту, яку позичальник зможе виплатити з урахуванням особливостей самого позичальника. Особливістю даної методики оцінки кредитоспроможності позичальника є оцінка платоспроможності позичальника з позиції погашення кредиту в строк [8].

Основними перевагами зазначених вище методів оцінки є достатня частка точності у визначеності кредитоспроможності потенційного позичальника, швидкість та можливість проведення експрес-аналізу заявки на кредит у присутності клієнта (скорингова оцінка та оцінка на основі фінансових показників платоспроможності) та вироблення до будь-якого потенційного позичальника індивідуального підходу (кредитна історія та андеррайтинг). Недоліками цих методів є зрівняння розрахункових кількісних показників кредитоспроможності клієнтів банку з нормативними, що виключає комплексний підхід до аналізу позичальника в наявних умовах (скорингова оцінка та оцінка на основі фінансових показників платоспроможності), та наявність чинника суб'єктивної оцінки працівником банку (кредитна історія та андеррайтинг).

Таким чином, аналіз підходів до оцінки кредитоспроможності клієнтів банку довів необхідність розроблення моделі, за допомогою реалізації якої можна проаналізувати всіх клієнтів банку та визначати нижню межу кредитоспроможності, що й буде нормативним показником, враховуючим наявні умови. Водночас базовим методом для моделі має бути найпоширеніший та найпростіший – скорингова оцінка. Інструментом, що дасть змогу проаналізувати велику кількість інформації, є метод інтелектуального аналізу даних.

Під інтелектуальним аналізом даних клієнтів банку слід розуміти процес виокремлення та аналізу великої кількості неструктурованої інформації з використанням статистичних та математичних методів для отримання нових, структурованих даних щодо кредитоспроможності клієнтів банку [2].

Виділяють такі типові етапи, які супроводжують рішення задач інтелектуального аналізу даних [1; 6; 10]:



Рис. 1. Етапи інтелектуального аналізу даних клієнтів банку

Інтелектуальний аналіз даних являє собою сукупність сучасних методів добування знань [7]. Науковий напрям, предметом вивчення якого є будь-яка інтелектуальна діяльність людини, що підкоряється заздалегідь відомим законам, є штучний інтелект.

Таким чином, використання методів штучного інтелекту під час моделювання оцінки кредитоспроможності клієнтів банку дасть змогу прискорити процес аналізу даних у процесі здійснення повторюваних та трудомістких задач.

Методика використання методу штучного інтелекту полягає у такій послідовності дій.

1. Визначення вихідних величин.

У моделі оцінки кредитоспроможності клієнтів банку передбачається використання 12 факторів, які було вибрано за допомогою аналізу факторів, що використовуються у скорингових моделях різноманітних банків: стать; наявність житла; наявність додаткових доходів; наявність автомобіля; наявність дітей; наявність вищої освіти; вік; співвідношення доходу до бажаного кредиту (обсяг доходу повинен бути у два рази більше обсягу кредиту); наявність сім'ї; стаж (до п'яти років та більше п'яти років); наявність гаранту за кредитом; наявність заборгованості за кредитами в інших банках.

Більшість інформації клієнт надає банку під час заповнення анкети на перевірку отримання кредиту, а інформацію про наявність заборгованості банк дізнається через бюро кредитних історій.

2. Перетворення категорійних даних у числові.

Для цього береться характеристика, що має категорійні дані, та розбивається на кілька

бінарних стовпців, замість одного стовпця, наприклад «стать», з'являється два: один для чоловіків, інший для жінок.

3. Здійснення прогнозу.

Всі розрахунки та прогноз здійснюється у ПП MS Excel, що значно спрощує використання моделі для працівників банку. Для здійснення прогнозу було використано лінійну регресію.

Формула для прогнозування є такою:

$$P = \sum_{i=1}^n \sum_{j=1}^m k_i * X_{ij}, \quad (1)$$

де k_i – коефіцієнт i -го фактору;

x_{ij} – дані за i -им фактором за j -им клієнтом.

Цільовою функцією виступає значення квадрата відхилень від значення стовпця, що характеризує кредитоспроможність до прогнозного значення:

$$Z = \sum_{j=1}^m P_j * Cred_j \rightarrow \min \quad (2)$$

Маємо такі обмеження:

$$\begin{cases} k_i \geq -1 \\ k_i \leq 1 \end{cases} \quad (3)$$

За допомогою вбудованого в ПП MS Excel інструменту «Пошук рішень» знаходяться коефіцієнти для кожного фактору лінійної моделі та мінімальна сума квадратів відхилень (цільова функція).

4. Оцінка отриманої моделі за допомогою коефіцієнту Фішера, R-квадрату та коефіцієнту Стьюдента.

5. Перевірка значущості кожного з факторів за допомогою побудови матриці суми квадратів та векторних добутків та її обчислення стандартного відхилення для кожного з коефіцієнтів, t -статистики та p -значення помилки.

На цьому етапі аналізу було визначено, що неістотними вважаються фактори «стать», «житло», «вік», «стаж». Отже, подальші дії будуть проводитися з такими факторами: «наявність дітей», «вища освіта», «дохід у два рази більше обсягу кредиту», «сімейний стан» та «наявність гаранту».

Інструмент «Пошук рішення» було запущено за новими даними, та на цій ітерації цільова функція знову дорівнює попередньому значенню. Отже, доцільно спробувати не лінійну регресію, а логістичну. Логістична регресія є типовим алгоритмом побудови моделі оцінки кредитоспроможності, який дає змогу обчислити рейтинг позичальника, зробити висновки, чи можна надавати цьому клієнту кредит, та отримати інтерпретацію результату у зручному вигляді.

Алгоритм використання логістичної регресії є аналогічним лінійній. Тобто необхідно знову повторити 1–5-й етапи.

Останнім кроком під час побудови логістичної моделі є переоптимізація моделі, тобто оптимум буде знаходитися не мінімізацією суми квадратів, а максимізацією значення логарифмічної функції правдоподібності для кожного прогнозу.

Таким чином, на основі реалізованої навчаючої моделі оцінки кредитоспроможності можна зробити висновок, що дана оцінка може бути проведена за допомогою лінійної та логістичної моделей, проте щоб визначити, яка модель є більш точною, необхідно провести інтерпретацію отриманих результатів.

Для оцінки якості моделей треба побудувати криву помилок для кожної з моделей та порівняти їх. Для початку треба зробити попередні розрахунки: мінімальне прогнозне значення, максимальне, границі діапазонів, точність, вибірковість, частка хибнопозитивних результатів та чуттєвість.

На основі даних, що було використано для реалізації моделі оцінки кредитоспроможності клієнтів банку, було побудовано криві помилок, що наведено на рис. 2.

Як можна побачити з рис. 2, моделі більш-менш схожі, проте найбільшу площу під кривою має лінійна модель. На даному рисунку можна побачити, що лінійна модель здатна ідентифікувати 19% кредитоспроможних клієнтів без жодної помилки, а логістична – 19,2%, та лінійна модель може досягти частки дійсно вірних результатів у 61% при 12,4% хибних спрацювань, а логістична – 56% при 8,8% хибних спрацювань. Порівнявши ці моделі, можна зробити висновок, що вони є більш-менш схо-

жими, проте логістична дає трохи більш точні результати.

Було визначено, що кредитоспроможним уважатиметься той клієнт, прогнозне значення якого, прораховане за допомогою логістичної моделі, перевищує 0,62. Клієнти, значення у яких буде перевищувати 0,87, будуть уважатися абсолютно кредитоспроможними з мінімальною часткою ризику.

Висновки з проведеного дослідження. На основі методу штучного інтелекту розроблено та реалізовано модель оцінки кредитоспроможності клієнтів банку. Визначено, що лінійна та логістична моделі надають схожі результати моделювання, проте логістична має точніші результати. Побудовано криві помилок для подальшого аналізу, та визначено, що лінійна модель здатна ідентифікувати 19% кредитоспроможних клієнтів без жодної помилки, а логістична – 19,2%. Визначено нижню межу кредитоспроможності на основі побудованих моделей – 0,62 та значення, за яким клієнт може вважатися абсолютно кредитоспроможним з мінімальною часткою ризику – 0,87. Перевагою запропонованої моделі оцінки кредитоспроможності клієнтів банку є доповнення методу скорингової оцінки, що широко використовується, методом штучного інтелекту, що дає змогу навчити наявну модель, побудовану в ПП MS Excel, визначати нижню межу кредитоспроможності. Реалізація моделі оцінки кредитоспроможності клієнтів банку істотно підвищує ефективність визначення неплатоспроможних клієнтів банку, оскільки під час визначення нормативних показників спирається на реальні характеристики клієнтів.

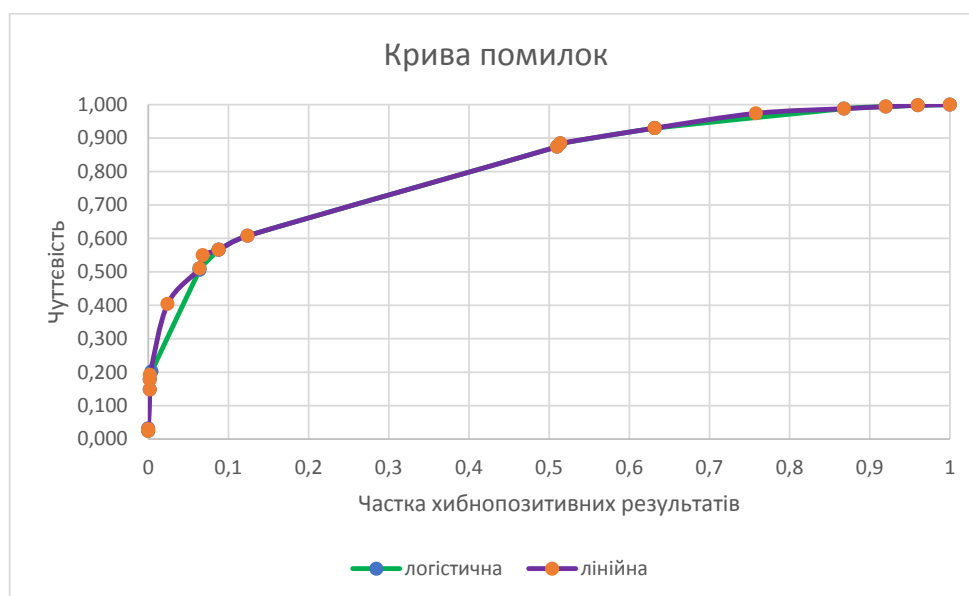


Рис. 2. Крива помилок для логістичної та лінійної моделей

Список використаних джерел:

1. Афанасьева С.В. Технология интеллектуального анализа данных: учеб. пособ.; Нац. исслед. ун-т «Высшая школа экономики», Санкт-Петербургский филиал. М.: Высшая школа экономики, 2013. 152 с.
2. Анализ данных и процессов: учеб. пособ. / А.А. Барсегян и др.; 3-е изд., перераб. и доп. СПб.: БХВ-Петербург, 2009. 512 с.
3. Вовк В.Я., Хмеленко О.В. Кредитування і контроль: навч. посіб. К.: Знання, 2008. 463 с.
4. Ворошилова И.В., Сурина И.В. К вопросу о совершенствовании механизма оценки кредитоспособности индивидуальных заемщиков. Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2005. № 16. С. 100–111.
5. Groshovo-kreditna ta finansova statistika – НБУ. URL: <https://bank.gov.ua/doccatalog/document?id=70285216>.
6. Замятин А.В. Интеллектуальный анализ данных : учеб. пособ. Томск: ТГУ, 2016. 120 с.
7. Кернасюк Ю.В. Інтелектуальний аналіз даних для навчання і практичного застосування в економіці. Наукові праці Кіровоградського національного технічного університету. Економічні науки. 2014. Вип. 26. С. 307–314.
8. Кургузкин А.Д., Просалова В.С. Андеррайтинг как способ оценки платежеспособности заемщика. Вестник Владивостокского государственного университета экономики и сервиса. 2012. № 1. С. 146–151.
9. Лаврушин О.И. Банковское дело: учебник; 2-е изд., перераб. и доп. М.: Финансы и статистика, 2005. 672 с.
10. Лиманова Н.И., Седов М.Н. Поиск реквизитов физических лиц на базах данных с использованием технологии Data Mining. Открытое образование. 2012. № 5. С. 42–47.

Bilenko D.V., Syvytska I.G., Telenkova D.G.

MODELING OF THE BANK CLIENTS CREDITABILITY ASSESSMENT BY A ARTIFICIAL INTELLIGENCE METHOD

The article examines the existing methods of credit assessing, their shortcomings and advantages. An analysis of approaches to assessing the bank client's creditworthiness is proved the need to develop a model, which implementation help to analyze all the bank clients and determine the lower limit of creditworthiness, which will be a normative indicator, considering the existing conditions. The basic method for the model that should be the most common and simplest is a scoring estimate. A tool that allows to analyze a large amount of information is the method of data mining.

Based on the analysis of shortcomings of existing methods, requirements for the model of credit rating of bank clients are formulated.

The method of modeling the bank's creditworthiness assessment with the help of artificial intelligence method is proposed. Approach consists in the identification and analysis of a large number of unstructured information using statistical and mathematical methods in order to obtain new, structured data on the creditworthiness of the bank's clients.

The method of using the artificial intelligence method is in this sequence of actions:

1. determination of output quantities;
2. converting categorical data into numeric;
3. implementation of the forecast;
4. estimation of the obtained model using Fisher's coefficient, R-square and Student's coefficient;
5. checking the significance of each factors by constructing the sum of squares matrix and vector products, calculating the standard deviation for each coefficients, t-statistics and p-value of the error.

The model is implemented on the real data and its advantages are defined, which consists in the ability to teach the existing model to determine the lower limit of credit based on really existing customer characteristics, which significantly increases the efficiency of determining the bank's insolvent clients.

Key words: bank customers creditworthiness, scoring, artificial intelligence method.