

УДК: 330.43, 336.77

DOI: <https://doi.org/10.32782/2224-6282/167-23>**Кравченко В. Г.**кандидат економічних наук, доцент,
Київський національний економічний університет імені Вадима Гетьмана
ORCID: <https://orcid.org/0000-0003-0113-7933>**Кравченко Т. В.**кандидат економічних наук, доцент,
Київський національний економічний університет імені Вадима Гетьмана
ORCID: <https://orcid.org/0000-0002-1506-3595>**Kravchenko Volodymyr, Kravchenko Tetiana**
National Economic University named after Vadym Hetman

ОЦІНЮВАННЯ КРЕДИТОСПРОМОЖНОСТІ ПОЗИЧАЛЬНИКІВ КОМЕРЦІЙНОГО БАНКУ

Зосереджено увагу на оцінюванні кредитоспроможності потенційних клієнтів банків. Розглянуто різні підходи до трактування сутності кредитоспроможності позичальника. Придільено увагу рейтинговому оцінюванню потенційних позичальників банків. На підставі цього запропоновано модель кредитного скорингу. Модель представлено у вигляді скорингової карти, яка будується на результатах оцінювання логістичної регресії. Одним з основних етапів побудови скорингової карти є перевірка на достовірність отриманої моделі. Для перевірки якості моделі були вибрані критерії: значення площі під ROC-кривою та індекс GINI. Отримано скорингову карту, яку в подальшому можна включити в автоматичний процес прийняття рішень щодо видачі кредитів банками.

Ключові слова: кредитоспроможність, логістична модель, кредитний скоринг, кваліфікаційна якість.

ASSESSMENT OF CREDIT CAPACITY OF POSITIVE IN COMMERCIAL BANK

In the financial and economic crisis context, which is characterized by an increase in the number of insolvent business entities of the credit market, there are an increase in the share of doubtful loans in loan portfolios of the banks and the issue of studying the assessment of the creditworthiness of borrowers of a commercial bank becomes especially relevant.

The article considered various approaches to the interpretation of the essence of the creditworthiness of borrowers and various methods for assessing the creditworthiness of potential customers of banks, such as regression models, neural networks, a classification tree, genetic algorithms, scoring cards, which are the main tools for data mining. Note that the different models can be applied at various stages of assessing a bank's credit risk.

The analysis of the assessment of the creditworthiness of borrowers by banks gave grounds to propose a credit-scoring model. The model is presented in the form of a scoring card, which is based on the results of evaluating the logistic regression. Scoring maps are constructed on the assumption that «the past reflects the future». Accordingly, based on data on previously opened loans and analyzing the available information, it is possible to predict the result (behavior) of future borrowers.

To create a scoring card, the following business process was considered: the manager of a partner store of the bank fills in the client's personal data, after which the form is sent to several banks for consideration. Banks review the application, obtain information from external sources, and make decisions. The bank, in turn, must quickly and efficiently assess the client and immediately indicate the agreed loan.

The scoring model in the form of a scoring card is based on the results of evaluating the logistic regression in the R-Studio software package. The quality of the investigated model was checked by the area under the ROC-curve and the GINI index. According to the results of the study, we can conclude that this model can be included in the automatic decision-making process on the issuance of loans, which will allow banks to 1) reduce the time for a bank to decide to issue a loan; 2) be sure that the loan will be repaid by the borrower; 3) reduce the bank's credit risks.

Keywords: creditworthiness, logistic model, credit scoring, qualification quality.

JEL classification: C01, C02, C13, C15, E51

Постановка проблеми. В умовах фінансово-економічної кризи, яка характеризується зростанням кількості неплатоспроможних суб'єктів господарювання кредитного ринку, що спричиняє зростання частки сумнівних кредитів у кредитних портфелях банків, особливої актуальності набуває питання вивчення оцінки кредитоспроможності позичальників комерційного банку. Це зумовлено тим, що значна частина прострочених кредитів виникає внаслідок недостатньо глибокого аналізу кредитоспроможності позичальника на етапі прийняття рішення щодо кредитування. Така ситуація спричинена відсутністю новітніх методик оцінки кредитоспроможності позичальника, які б чітко

відповідали вимогам сьогодення. Тому одним з основних завдань для банків є здатність ними здійснювати кредитну діяльність та знижувати ризики, пов'язані з нею.

Аналіз останніх досліджень і публікацій. Питанню вивчення оцінки кредитоспроможності позичальників та пов'язаних із ними ризиків присвячено багато вітчизняних та зарубіжних праць, зокрема: Е. Брігхем, Л. Гапенські, П. Роуз, Дж.Ф. Сінкі, В.В. Вітлінського, О.В. Дзюблюка, Л.О. Примостки, Т.О. Терещенко тощо. Вони вивчали модернізацію методів та моделей оцінки кредитоспроможності боржника комерційними банками. Проте проблему оцінки кредитоспромож-

ності позичальників не можна вважати повністю розкритою, тому що наразі немає єдиної системи оцінки, яка б повністю описувала загальну картину кредитоспроможності позичальника.

Мета статті полягає у висвітленні існуючих підходів до оцінки кредитоспроможності позичальників та побудові моделі кредитного скорингу для визначення кредитного ризику банків.

Для досягнення поставленої мети були визначені такі завдання: 1) розглянути сутність та необхідність оцінки кредитоспроможності позичальника банку; 2) здійснити аналіз підходів до оцінки кредитоспроможності позичальника; 3) провести оцінку кредитоспроможності позичальника за допомогою моделі кредитного скорингу; 4) оцінити якість моделі за ROC-кривою та індексом GINI.

Виклад основного матеріалу. Проведення активних операцій банками повинно здійснюватися відповідно до умов, визначених у чинному законодавстві. В обов'язковому порядку кредитним комітетом здійснюється оцінка кредитоспроможності позичальника, ефективність проведення якої безпосередньо впливає на кредитну діяльність банку. Фінансово-кредитна установа зацікавлена в найбільш реальному відображенні стану кредитоспроможності позичальника з огляду на те, що чим більш ефективно буде здійснена оцінка, тим менша ймовірність настання кредитного ризику, а отже, тим більшою є вірогідність погашення позичальником своїх зобов'язань на користь банку.

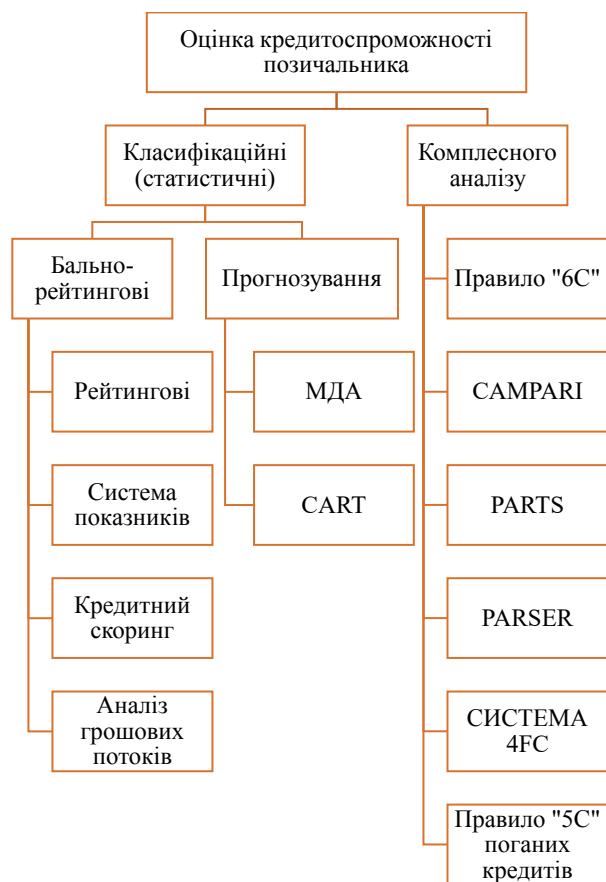


Рис. 1. Методи оцінки кредитоспроможності позичальника

Джерело: складено за [1]

Кредитоспроможність – це такий фінансовий стан, за якого позичальник здатен накопичити визначену суму коштів для погашення зобов'язань у визначений договором термін.

Зауважимо, що кредитоспроможність складається із сукупності елементів, висвітлення яких дає чітке уявлення про сутність досліджуваної економічної категорії, а саме: довіри банку до позичальника; наявності передумов отримання кредиту (кредитна історія); аналізу становища позичальника на ринку; репутації позичальника; гарантій; ділової активності, тобто підтримки фінансової стійкості, платоспроможності.

Існування різних підходів до трактування сутності кредитоспроможності позичальника, відмінності методик аналізу, моделей оцінки та системи пояснює їх різноманітність у банківській системі, тому комерційний банк в Україні може самостійно вибирати найбільш оптимальні показники для оцінки та аналізу кредитоспроможності.

Схематично всі існуючі методики можна представити у вигляді класифікаційних (статистичних) методів оцінки та комплексного аналізу кредитоспроможності (рис. 1).

У нашій роботі зупинимось на моделюванні кредитних ризиків за допомогою кредитного скорингу.

Кредитний скоринг (від англ. score – бал, оцінка) є математичною або статистичною моделлю, за допомогою якої на основі кредитної історії попередніх клієнтів банк намагається визначити, наскільки велика ймовірність того, що конкретний потенційний позичальник поверне кредит у визначений термін [2, с. 88].

Процедура кредитного скорингу полягає у тому, щоб оптимізувати надання кредитів позичальникам. У цій процедурі використовується бальна система (ураховуючи вагові коефіцієнти) та відбувається класифікація потенційних позичальників з урахуванням рейтингу їхньої кредитоспроможності. Тобто кожному досліджуваному параметру, який описує кредитора-позичальника, присвоюється реальна оцінка в балах. На підставі статистичних досліджень даних кожному показнику встановлюється певна кількість балів – скорингова оцінка. Чим вище кредитоспроможність і сумлінність за тим або іншим показником, тим вищий бал. Тому основна ідея скорингу – визначення тих характеристик, які пов'язані зі ступенем надійності та платоспроможності потенційного позичальника.

Різними банками для побудови скорингової моделі використовуються різні методи, наприклад: моделі регресії, нейронні мережі, дерево класифікацій, генетичні алгоритми, скорингові карти, що є основними інструментами інтелектуального аналізу даних. Ці моделі можна використовувати на різних етапах оцінювання кредитного ризику банку.

Найбільш традиційною у банківській системі є лінійна регресійна модель множинної регресії [3, с. 34; 4, с. 237]:

$$p = w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n, \quad (1)$$

де p – ймовірність дефолту;

w – вагові коефіцієнти;

x – характеристики клієнта.

Проте недоліком моделі (1) є те, що в лівій частині рівняння знаходиться ймовірність, яка приймає зна-

чення від 0 до 1, а змінні в правій частині можуть приймати будь-які значення від $-\infty$ до $+\infty$.

Логістична регресія дає змогу подолати цей недолік:

$$\text{Log} \left(\frac{p}{1-p} \right) = w_0 + w_1 x_1 + w_2 x_2 + \dots + w_n x_n. \quad (2)$$

Ця модель потребує більш складніших розрахунків для отримання вагових коефіцієнтів, та обов'язковою умовою є те, що в ній не повинно бути сильно корельованих незалежних змінних. Проте слід зазначити, що модель (2) є лідером серед скорингових (рейтингових) систем. Основною перевагою моделі (2) є те, що клієнтів можна розділити на дві групи (1 – потенційно хороший клієнт, 0 – поганий) і на декілька груп ризику. Останнім кроком є те, що модель (2) потрібно перевірити на достовірність та апробувати на реальних даних. Модель повинна давати коректні та якісні прогнози на практиці.

Дерево класифікації і нейроні мережі являють собою системи, які поділяють клієнтів на групи, у середині яких рівень ризику приблизно однаковий і максимально відрізняється від рівня ризику інших груп [4, с. 238].

Ще одним із методів кредитного скорингу є скорингові карти.

У нашій роботі за допомогою кредитного скорингу дослідимо кредитні ризики банків із реальної автоматичної системи прийняття рішень щодо видачі товарного кредиту. Розглядається такий процес системи: менеджер магазину – партнера банку заповнює анкетні дані клієнта, після чого аплікаційна інформація надсилається на розгляд у декілька банків. Банки розглядають заявку, отримують інформацію із зовнішніх джерел та приймають рішення. Банк, своєю чергою, повинен швидко й якісно оцінити клієнта та відразу вказати, чи погоджено кредит.

Процес розроблення скорингової моделі може відрізнятися залежно від вибраної предметної галузі, експертних оцінок аналітика-розробника, кінцевої мети та інших чинників.

Структуру створення та впровадження всього процесу в системі R-Studio наведено на рис. 2.

Аналіз та збирання даних. На цьому етапі отримано клієнтські дані для роботи, які було структуровано за встановленим зразком звіту щодо товарних заявок у таблиці Excel. Вибір предикторів зумовлений наявними даними, такими як: технічні дані щодо заявки; аплікаційні дані, надані клієнтом; дані щодо вибраного продукту, партнера, щодо клієнтів із зовнішніх сервісів.

Для дослідження та побудови логістичної моделі взято 86 факторних ознак та 8 412 спостережень. Дані містять такі факторні ознаки, як: стать, вік, паспорт, адреса, сімейний стан, освіта, діти, посада тощо.

Робота з даними виконувалася у зручному середовищі R-Studio, яке працює на основі мови програмування R. Слід зазначити, що починати роботу із середовищем можна, коли дані очищені та структуровані, для цього достатньо мати попередньо підготовлений Excel-файл. Фрагмент даних представлено на рис. 3.

Скорингові карти будуються виходячи з припущення про те, що «минуле відображає майбутнє». Відповідно, базуючись на даних щодо відкритих раніше кредитів і аналізуючи наявну інформацію, можна передбачити результат (поведінку) майбутніх позичальників. Для того щоб коректно виконати цей аналіз, потрібно зібрати необхідні дані за певний проміжок часу, а потім здійснити їх моніторинг протягом іншого визначеного відрізка часу й оцінити. Зібрані дані (змінні) поряд із відповідною класифікацією (мета: «хороший»/«поганий») становлять основу для розроблення скорингової карти.

«Вікно дозрівання» – це проміжок часу, коли позичальник мав можливість себе проявити (наприклад, мета: 90+). «Вікно вибірки» – це проміжок часу, коли ті чи інші позичальники відбираються для аналізу (потрапляють у вибірку). Рекомендується також проаналізувати, який період «дозрівання» є ідеальним для того чи іншого продукту (регіону, типу клієнта тощо). У деяких випадках, таких як шахрайство і банкрутство, часовий період уже відомий або визначений.

Припустимо, що термін кредитування менше шести місяців і вибірка ще «не визріла», визначимо «поганих» клієнтів – цільову змінну – за причиною відмови. Увійшли: відмова за негативною кредитною історією, за наявністю у розшуку, загублені паспорти, критичні групи фроду.

Із моделі також були виключені неякісні дані або нетипові випадки, які можуть змістити розподіл реальних значень. Також у цій моделі дані очистили від викидів шляхом зміни значень або вилучення запису. На рис. 4 представлено часовий ряд «Вік» позичальника.

Як можемо побачити (рис. 4), у межах 2 100 значення є викид зі значенням 4. Перевіряємо за датою народження і бачимо помилку – виправляємо її.

Одномірний аналіз. Розподіл характеристик. На даному етапі перевіряємо, які є дані для побудови моделі, їх тип, кількість та якість, визначаємо вагомість атрибутів. Важливим є визначення міри прогностичної здатності атрибута і характеристики – вагомість Weight of Evidence (WoE). Ця міра означає міру ймовірності «хорошого» результату в даній характеристиці, відношення частки «хороших» в атрибуті до частки «поганих» результатів. Після цього визначимо оцінку ступеня взаємозв'язку між незалежними змінними і залежною – показник інформаційного значення, або

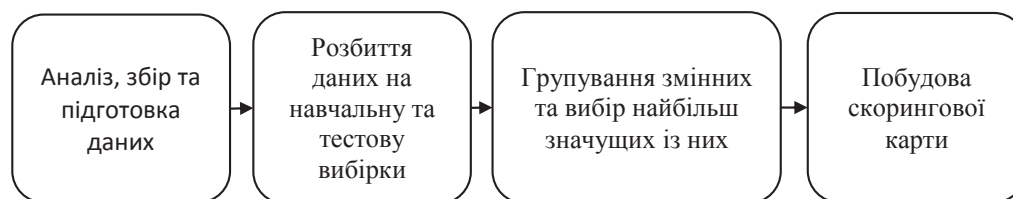


Рис. 2. Процес створення скорингової карти в R-Studio

Джерело: складено за [6, с. 29]

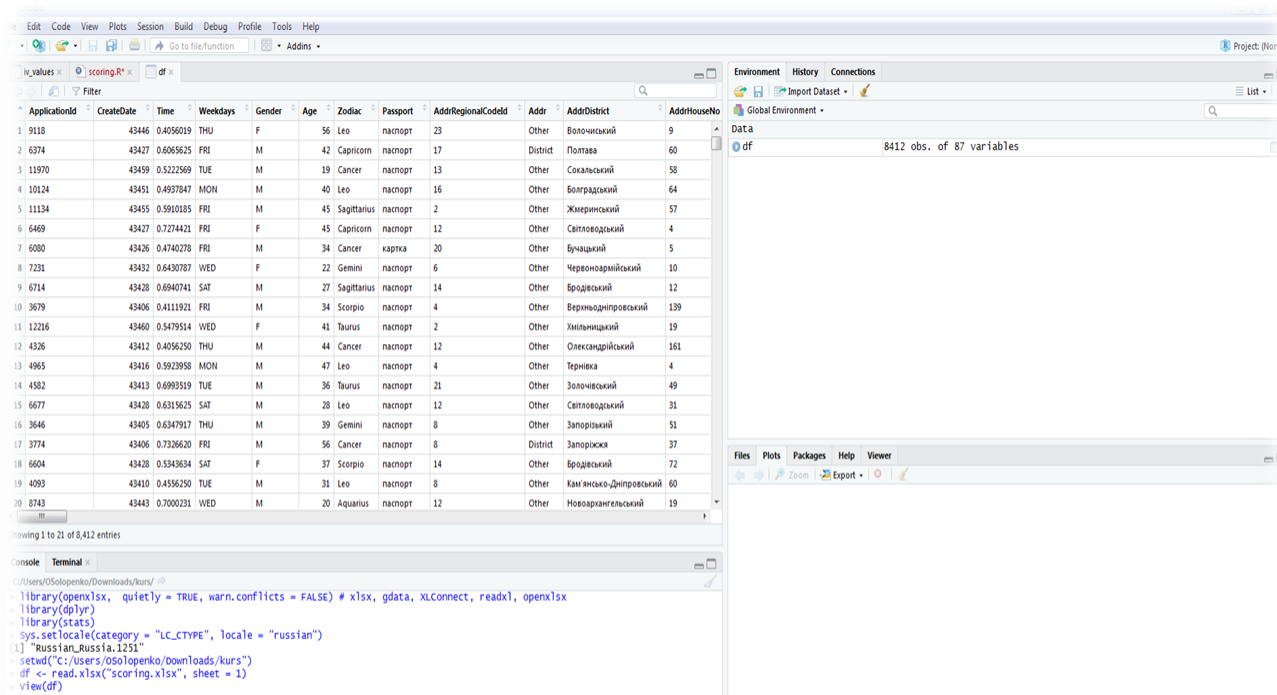


Рис. 3. Фрагмент генеральної сукупності для побудови моделі в R-Studio

Джерело: розроблено авторами

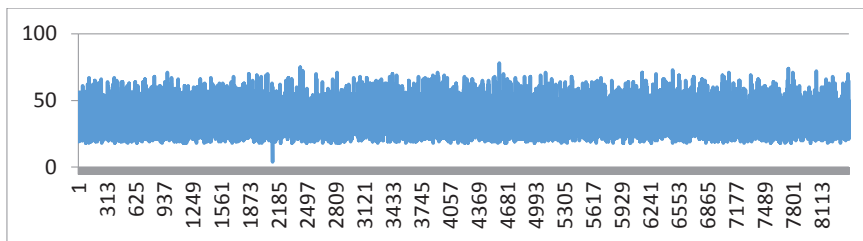


Рис. 4. Графік числового ряду змінної «Вік» позичальників

Джерело: розроблено авторами

IV (Information value) «предиктивна сила» характеристики, за формулою:

$$IV = \sum_1^n (d_i^{(1)} - d_i^{(2)}) \cdot \ln \left(\frac{d_i^{(1)}}{d_i^{(2)}} \right), \quad (3)$$

де $d_i^{(1)}$ і $d_i^{(2)}$ – відносна частка поганих та хороших клієнтів, n – кількість категорій незалежних змінних [6, с. 25]. Чим вищий цей показник, тим більшу вагу має значення змінної під час побудови моделі (табл. 1):

Таблиця 1

Значущість незалежної змінної за показником IV

Information Value	Предиктивна сила
< 0.02	Марне для прогнозування
0.02 to 0.1	Слабкий предиктор
0.1 to 0.3	Середній предиктор
0.3 to 0.5	Сильний предиктор
>0.5	Підозрілий предиктор

Джерело: складено за [6, с. 25]

Якою б сильною не була характеристика, необхідно брати до уваги бізнес-логіку: чи дійсно дана залеж-

ність має «економічну» природу? чи не викликана дана залежність впливом інших чинників? чи немає помилки в розрахунках? Великі IV характеристики завжди підозрілі.

Таблиця 2

Інформація про значущість незалежної змінної до групування

Variable	info_value
PartnerStoreCode	0,524697
CH	0,478453
JurDistrict	0,416215
AddrDistrict	0,391286
.....
Addr	0,000248

Джерело: розроблено авторами

У табл. 2 бачимо, що найбільший вплив мають конкретний магазин, в якому клієнт отримував кредит, групова оцінка його кредитної історії та адреса прописки.

Заключним моментом одномірного аналізу скорингових показників є групування всередині змінних на

Таблиця 3

Фрагмент групування всередині змінних на класи

variable	bin	count	good	bad	badprob	woe	bin_iv	total_iv	breaks
Time	[-Inf,0.41)	300	242	58	0,1933	-0,2577	0,0033	0,0126	0.41
Time	[0.41,0.51)	1738	1331	407	0,2342	-0,0141	0,0001	0,0126	0.51
...
Time	[0.73, Inf)	383	297	86	0,2245	-0,0686	0,0003	0,0126	Inf

Джерело: розроблено авторами

класи. Особливо це стосується інтервальних і якісних змінних із великою кількістю характеристик, наприклад професія. Групування всередині показників дає змогу легше зрозуміти існуючі залежності в моделі, скорингові бали стають більш прозорими і стійкими до невеликих змін у клієнтській базі, вирішує проблему екстремальних величин і рідкісних значень і надає більше статистичної значущості скоринговими змінним.

Інформацію щодо кожного предиктора отримали у вигляді табл. 3.

Із табл. 3 бачимо, що перша група має період від 0 до 0,41 та вміщує 300 значень, з яких 58 – «поганих». Ймовірність таких значень у цій групі – 0,1933, Woe –0,2577. Предиктивна сила всього Time – 0,0126, а конкретного розглянутого значення – 0,0033.

Знову обрахуємо IV вже після групування:

Таблиця 4

Інформація про значущість незалежної змінної після групування

variable	info_value
CH_woe	0,315045
FotoVerification_woe	0,205543
Age_woe	0,130567
.....
Addr_woe	0,000248

Джерело: розроблено авторами

У табл. 4 бачимо, що після групування трійка лідерів змінилася (тепер найбільший вплив здійснюють кредитна історія, фотоверифікація та вік), а значення значущості здебільшого зменшилися.

Результати групування представлено на рис. 5.

Дослідження кореляції. Маємо оброблені дані, інформацію стосовно предикторної сили характеристик та мету – дослідити зв'язок між даними боржника та їхнім впливом на ймовірність, тобто тепер можемо створити модель скорингової оцінки кредитоспроможності клієнта.

Усі змінні, що використовуються в моделі, необхідно перевірити на наявність кореляції між собою для коректнішого відображення впливу кожного з критеріїв на результат (рис. 6).

Із рис. 6 бачимо, що, наприклад, AddrRegionalCodeId корелює з JurRegionalCodeId, що не дивно, адже зазвичай адреса прописки та адреса проживання вказуються однакові, тому видаляємо корелюючі характеристики.

Багатомірний аналіз (побудова скорингової карти). На основі логістичної моделі було створено вибірку даних щодо оцінки кредитоспроможності клієнтів. Залежна змінна Y інтерпретується як ймовірність дефолту позичальника, це бінарна змінна, яка приймає значення 1, якщо дефолт, і значення 0 – у протилежному випадку. Шляхом перетворень отримаємо фінальний вибір предикторів для моделі. Чим більше зірочок, тим впливовішим він є. Предиктори, яких немає в списку, були відкинуті (табл. 5).

Для оцінки якості класифікації використаємо аналіз ROC-кривих (від англ. Receiver Operator Characteristic), що дасть змогу вибрати оптимальний поріг ймовірності поділу «поганих» і «хороших» позичальників для досягнення прийнятного рівня чутливості і специфічності моделі [5, с. 9–10]. Для аналізу класифікаційної здатності скорингової моделі використовують показник площі під ROC-кривою – AUC (від англ. Area

До групування

variable	info_value
PartnerStoreCode	0,524697
CH	0,478453
JurDistrict	0,416215
AddrDistrict	0,391286
RequestedLoanAmount	0,271082
PurchasePrice	0,27092
Time	0,258122
Age	0,255588
WorkExperience	0,241793
Vnesok_vid	0,193582
Vnesok_kil	0,187628
FinancialIncomes0	0,150589
FotoVerification	0,128882

Після групування

variable	info_value
CH_woe	0,315045
FotoVerification_woe	0,205543
Age_woe	0,130567
Vnesok_vid_woe	0,109097
Vnesok_kil_woe	0,102942
PartnerStoreCode_woe	0,100462
PartnerName_woe	0,06757
WorkExperience_woe	0,049631
AddrDistrict_woe	0,049224
JurDistrict_woe	0,045915
LoanProductName_woe	0,043627
JurRegionalCodeId_woe	0,043527
AddrRegionalCodeId_woe	0,043324

Рис. 5. Групування змінних

Джерело: розроблено авторами

gb	Time_woe	Gender	Age_woe	Zodiac	Passport	AddrReg	Addr_woe	AddrDist	AddrFlat	JurAddr	JurJurDist	JurFlat	FamilySt	Educatio	WorkExp	WorkPos	WorkPlac	ContactP	ContactF	Financia	Financi	PartnerN	PartnerS	Purchase	LoanProd	Request	Vnesok	Vnesok	Ch_woe	FotoVerif				
gb	1	0.11854	0.05124	0.03744	0.15384	0.04413	0.04883	0.08872	0.00666	0.00996	0.03084	0.02846	0.11183	0.08671	0.02489	0.05024	0.04445	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	
Time_woe	0.11854	-0.0001	-0.00173	0.02324	0.03393	0.04883	0.08872	0.00666	0.00996	0.03084	0.02846	0.11183	0.08671	0.02489	0.05024	0.04445	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485		
Gender	0.05124	-0.0001	-0.01273	0.02324	0.03393	0.04883	0.08872	0.00666	0.00996	0.03084	0.02846	0.11183	0.08671	0.02489	0.05024	0.04445	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485		
Age_woe	0.03744	-0.01273	-0.01853	1	0.01508	0.00487	0.01106	-0.05817	0.01124	0.0062	0.04733	0.05078	0.01434	-0.01601	0.07839	0.01529	0.08867	0.08564	0.06292	0.03859	0.19281	0.05782	0.23072	0.04839	-0.01276	-0.01452	-0.02106	0.03238	0.03311	-0.02793	0.00381	-0.04024	-0.04073	
Zodiac	0.15384	0.02324	0.03393	0.01508	1	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989	-0.00889	-0.00889	-0.07748	0.03745	-0.02471	0.01715	-0.00683	0.03182	0.05312	-0.00164	-0.04738	0.02511	0.00272	-0.02933	0.03004	-0.00747	-0.0133	0.30027	-0.2493	-0.04073		
Passport	0.04413	0.00487	0.01106	-0.00027	-0.16094	1	-0.00027	-0.00376	0.00777	-0.00274	0.00337	0.01969	0.00677	0.01388	0.00444	0.03261	0.01001	0.0133	0.00299	0.02784	-0.00299	-0.00249	-0.00106	-0.003	-0.01879	0.00663	0.01139	-0.00376	0.00671	0.00919	0.01043	0.00278	0.00296	-0.01311
AddrReg	0.08872	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	1	-0.00027	-0.00376	0.00777	-0.00274	0.00337	0.01969	0.00677	0.01388	0.01001	0.0133	0.00299	0.02784	-0.00299	-0.00249	-0.00106	-0.003	-0.01879	0.00663	0.01139	-0.00376	0.00671	0.00919	0.01043	0.00278	0.00296	-0.01311
Addr_woe	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	1	-0.00027	-0.00376	0.00777	-0.00274	0.00337	0.01969	0.00677	0.01388	0.01001	0.0133	0.00299	0.02784	-0.00299	-0.00249	-0.00106	-0.003	-0.01879	0.00663	0.01139	-0.00376	0.00671	0.00919	0.01043	0.00278	0.00296	-0.01311	
AddrDist	0.03084	0.02846	0.11183	0.08671	0.02489	0.05024	0.04445	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989
AddrFlat	0.02846	0.11183	0.08671	0.02489	0.05024	0.04445	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989	
JurAddr	0.11183	0.08671	0.02489	0.05024	0.04445	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989		
JurJurDist	0.08671	0.02489	0.05024	0.04445	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989			
JurFlat	0.02489	0.05024	0.04445	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989				
FamilySt	0.02489	0.05024	0.04445	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989				
Educatio	0.05024	0.04445	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989					
WorkExp	0.04445	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989						
WorkPos	0.05688	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989							
WorkPlac	0.05815	0.07933	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989								
ContactP	0.02326	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989										
ContactF	0.08115	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989											
Financia	0.06635	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989												
Financi	0.05	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989													
PartnerN	0.02625	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989														
PartnerS	0.13008	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989															
Purchase	0.03624	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989																
LoanProd	0.0979	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989																	
Request	0.13223	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989																		
Vnesok	0.12796	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989																			
Vnesok	0.1317	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989																				
Ch_woe	0.23313	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989																					
FotoVerif	0.21485	0.00666	0.00996	0.00487	0.01106	-0.00027	-0.16094	0.00559	0.02002	0.01422	-0.00955	-0.00989																						

Рис. 6. Кореляційна матриця

Джерело: розроблено авторами

Таблиця 5

Оцінки параметрів логістичної моделі

Coefficients:	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.10695	0.04655	-23.777	< 2e-16 ***
Time_woe	1.15359	0.16221	7.112	1.15e-12 ***
Weekdays_woe	0.63825	0.36924	1.729	0.083895 .
Age_woe	0.83700	0.12755	6.562	5.30e-11 ***
Zodiac_woe	1.17413	0.43028	2.729	0.006357 **
Passport_woe	1.63777	0.40064	4.088	4.35e-05 ***
JurRegionalCodeId_woe	0.65340	0.17393	3.757	0.000172 ***
JurDistrict_woe	0.50815	0.22106	2.299	0.021524 *
FamilyStatus_woe	0.80824	0.25460	3.175	0.001501 **
EducationLevelName_woe	0.96436	0.38402	2.511	0.012031 *
workPositionTypeName_woe	-0.85030	0.45874	-1.854	0.063807 .
WorkExperience_woe	0.78795	0.18343	4.296	1.74e-05 ***
WorkPlaceAreaName_woe	0.63019	0.26171	2.408	0.016042 *
ContactPersonsTypeId0_woe	0.63670	0.19008	3.350	0.000809 ***
ContactPersonsTypeId1_woe	1.02051	0.23538	4.336	

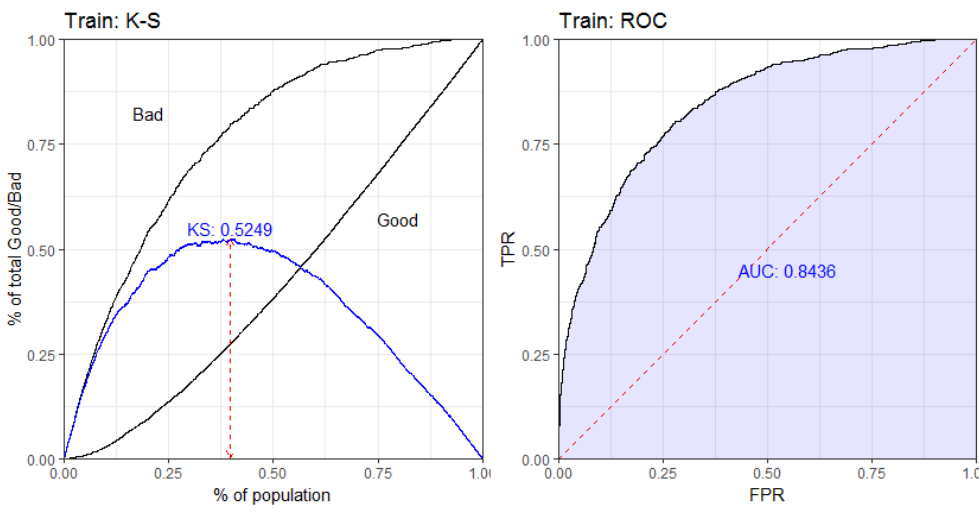


Рис. 7. Статистика тренувальної вибірки

Джерело: розроблено авторами

Розрахований коефіцієнт Джині становив 0,6872, що є досить хорошим показником.

Якість класифікації розробленої на основі логістичної регресії скорингової карти можна оцінити, проаналізувавши розподіл скорингових балів позичальників. Під час ранжирування позичальників за хорошою скоринговою картою «погані» і «хороші» позичальники повинні мати максимально відмінні скорингові бали. Наприклад, «поганим» позичальникам присвоювався менший бал, аніж «хорошим». У результаті група «поганих» кредитів повинна мати в сумі менше балів, аніж група «хороших». Чим більш явно будуть розділені розподілу скорингових балів для «поганих» і «хороших» кредитів, тим ефективніше буде працювати скорингова карта. Ідеальною вважається скорингова карта, за якою розподіл балів у «поганих» і «хороших» позичальників не перетинається, а знаходиться поруч один з одним.

Розраховуємо саму скорингову карту з балами, які програмний комплекс записує поряд із кожним значенням предиктора, створюючи тим самим довідник (табл. 6):

Також подивимося, як розподіляться бали на тренувальній вибірці, яка розраховує кредитний бал, урахувавши попередні розрахунки (табл. 7).

У табл. 6 та 7 бачимо, звідки були взяті сумарні бали та як були розставлені бали за кожне значення фінальних предикторів декількох із позичальників.

За розподілу Гауса (рис. 8) бачимо, що більшість «поганих» позичальників зосереджено в лівому боці графіка, адже чим менший бал, тим більша приналежність до «поганих», а це означає, що такий розподіл є досить хорошим.

Однак «хороші» позичальники не зосереджені у правій частині графіку, де мали б знаходитися. Багато «хороших» позичальників отримали низький скоринговий бал, що показує недоліки моделі. Результатом моделювання є число, конкретний скоринговий бал, який обраховується окремо для кожного позичальника.

Таким чином, ми можемо встановити пороговий бал відмови позичальникам, які отримали менше 826 балів, та бал для затвердження без перевірки вище 1 063. Відповідно, із 826 до 1063 балів виникне сіра зона, яку потрібно буде перевіряти додатково та поступово вносити зміни в модель.

Висновки. Проведене дослідження сучасних моделей оцінювання кредитоспроможності позичальників свідчить про їх недосконалість та існування низки

Таблиця 6

Фрагмент довідника скорингової карти з балами для окремих значень предиктора

variable	bin	count	count_dis	good	bad	badprob	woe	bin_iv	total_iv	breaks	is_special	points
Time	[-Inf,0.386	83	0,014564	64	19	0,228916	-0,04363	2,74E-05	0,080833	0.386	ЛОЖЬ	26
Time	[0.386,0.391	40	0,007019	35	5	0,125	-0,7751	0,003364	0,080833	0.391	ЛОЖЬ	87
Time	[0.391,0.396	43	0,007545	32	11	0,255814	0,102972	8,22E-05	0,080833	0.396	ЛОЖЬ	14
Time	[0.396,0.401	36	0,006317	29	7	0,194444	-0,25057	0,00037	0,080833	0.401	ЛОЖЬ	43

Джерело: розроблено авторами

Таблиця 7

Фрагмент розподілення балів на тренувальній вибірці

Time_poi	Weekday	Age_poi	Zodiac_pc	Passport_JurRegion	JurDistrict	FamilySta	Education	WorkPosi	WorkExpe	WorkPlaci	ContactPe	ContactPe	FinancialI	PartnerNe	PurchaseF	Requeste	Vnesok_v	CH_points	FotoVerif	score	
51	18	52	12	26	30	23	31	19	6	44	34	34	15	31	43	11	33	9	-11	79	590
28	18	3	30	26	26	20	31	19	21	44	33	27	15	0	10	42	37	17	287	124	858
19	18	40	30	26	8	20	33	19	18	25	16	35	65	30	38	27	20	17	-2	-62	440
32	18	23	26	-24	47	44	31	15	18	50	17	27	15	18	10	4	25	37	-2	81	512

Джерело: розроблено авторами

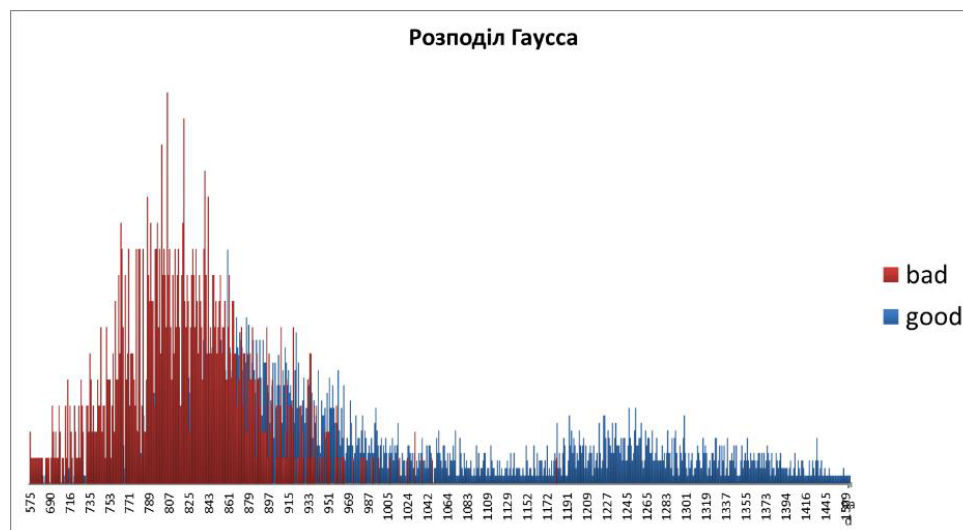


Рис. 8.

Джерело: розроблено авторами

недоліків. Класифікаційні методи оцінки кредитоспроможності позичальника недостатньо уваги приділяють якісним показникам його діяльності, не досліджується вплив зовнішнього середовища на здатність підприємства обслуговувати кредитні кошти.

Тому в роботі був розглянутий приклад скорингової моделі, який дає можливість банкам: 1) скоротити час на прийняття банком рішення на видачу кредиту; 2) бути впевненими в тому, що кредит буде повернений позичальником; 3) знизити кредитні ризики банку.

Для розв'язання поставленої задачі було вибрано скорингову модель у вигляді скорингової карти, яка будується на результатах оцінювання логістичної регресії в програмному комплексі R-Studio. За значен-

нями площі під ROC-кривою та індексом GINI була перевірена якість скорингової моделі. Отримано скорингову карту, яку можна включати в автоматичний процес прийняття рішень щодо видачі кредитів.

Напрямами подальших досліджень є вдосконалення розглянутої методики скорингового аналізу та визначення переваг і недоліків упровадження автоматизованих скорингових систем у роботу комерційних банків. Перспективним напрямом досліджень є також використання комбінованих процедур прогнозування кредитоспроможності клієнтів на основі альтернативних, ідеологічно різних, методів аналізу даних, що забезпечують значне підвищення якості оцінок прогнозів.

Список використаних джерел:

1. Виговський В. Г. Методичні підходи до оцінки кредитоспроможності суб'єктів господарювання комерційними банками. *Облік і фінанси АПК*. URL: <http://magazine.faaf.org.ua/metodichni-pidhodi-do-ocinki-kreditospromozhnosti-sub-ektiv-gospodaryuvannya-komerciyrimi-bankami.html> (дата звернення: 12.02.2021).
2. Стечишин Т.Б. Сучасні банківські методики визначення кредитоспроможності позичальника – фізичної особи. *Наука молода*. 2015. № 23. С. 82–93. URL: <http://dSPACE.tneU.edu.ua/handle/316497/3433> (дата звернення: 16.02.2021).
3. Данилов В.Я., Жиров О.Л., Бідюк П.І. Оцінювання кредитних ризиків методами інтелектуального аналізу даних. *Системні дослідження та інформаційні технології*. 2017. № 1. С. 33–48. URL: <https://core.ac.uk/download/pdf/84839117.pdf> (дата звернення: 16.02.2021).
4. Яблоков А.І. Рейтингове моделювання банківських ризиків. *Економіко-математичне моделювання соціально-економічних систем*. 2009. № 14. С. 229–245. URL: <http://dSPACE.nbuv.gov.ua/bitstream/handle/123456789/46082/14-Yablokov2.pdf?sequence=1> (дата звернення: 18.02.2021).
5. Сорокін А.С. К вопросу валидации модели логистической регрессии в кредитном скоринге. *Науковедение*. 2014. № 2(21). С. 1–14. URL: <http://znanium.com/catalog/product/487886> (дата звернення: 18.02.2021).
6. Бакун С.А., Бідюк П.І. Методика побудови скорингових карт із використанням платформ SAS. *Наукові вісті Національного технічного університету України «Київський політехнічний інститут»*. 2016. № 2. С. 23–32. URL: https://ela.kpi.ua/bitstream/123456789/18170/1/5_Bakun.pdf.pdf (дата звернення: 10.02.2021).

References:

1. Vyhovskiy V.H. Metodichni pidkhody do otsinky kredyitospromozhnosti subiektiv hospodariuvannya komertsiiynymi bankamy [Methodical approaches to assessing the creditworthiness of commercial banks]. *Accounting and finance of agro-industrial complex (electronic journal)*. Available at: <http://magazine.faaf.org.ua/metodichni-pidhodi-do-ocinki-kreditospromozhnosti-sub-ektiv-gospodaryuvannya-komerciyrimi-bankami.html> (accessed 02 February 2021).
2. Stechyshyn T.B. (2015) Suchasni bankivski metodyky vyznachennia kredyitospromozhnosti pozychalnyka – fizychnoi osoby [Modern banking methods for determining the creditworthiness of the borrower – an individual]. *Nauka moloda*, vol. 23, pp. 82–93. Available at: <http://dSPACE.tneU.edu.ua/handle/316497/3433> (accessed 16 February 2021).

3. Danylov V.Ya., Zhyrov O.L., Bidiuk P.I. (2017) Otsiniuvannia kredytnykh ryzykiv metodamy intelektualnoho analizu danykh [Credit risk assessment by data mining methods]. *Systemni doslidzhennia ta informatsiini tekhnologii (international scientific and technical journal)*, vol. 1, pp. 33–48. Available at: <https://core.ac.uk/download/pdf/84839117.pdf> (accessed 16 February 2021).
4. Iablokov A.I. (2009) Reitynhove modeliuвання bankivskykh ryzykiv [Rating modeling of banking risks]. *Ekonomiko-matematychne modeliuвання sotsialno-ekonomichnykh system (Zbirnyk naukovykh prats MNNTs ItiS)*, vol. 14, pp. 229–245. Available at: <http://dspace.nbuv.gov.ua/bitstream/handle/123456789/46082/14-Yablokov2.pdf?sequence=1> (accessed 18 February 2021).
5. Sorokin A.C. (2014) K voprosu validacii modeli logisticheskoy regressii v kreditnom skoringe [On the issue of validating the logistic regression model in credit scoring]. *Naukovedenie (electronic journal)*, vol. 2, no. 21, pp. 1–14. Available at: <http://znanium.com/catalog/product/487886> (accessed 18 February 2021).
6. Bakun S.A., Bidiuk P.I. (2016) Metodyka pobudovy skorynhovykh kart iz vykorystanniam platform SAS [Methods of constructing scoring maps using SAS platforms]. *Naukovi visti Natsionalnoho tekhnichnoho universytetu Ukrainy «Kyivskiy politekhnichnyi instytut»*, vol. 2, pp. 23–32. Available at: https://ela.kpi.ua/bitstream/123456789/18170/1/5_Bakun.pdf.pdf (accessed 10 February 2021).