

РОЗДІЛ 5. РОЗВИТОК ПРОДУКТИВНИХ СИЛ І РЕГІОНАЛЬНА ЕКОНОМІКА

МОДЕЛЮВАННЯ КРЕДИТОСПРОМОЖНОСТІ КЛІЄНТА MODELLING CLIENT'S CREDITABILITY

УДК 519.86:336.77

<https://doi.org/10.32843/infrastruct39-44>

Аверкина М.Ф.

д.е.н., професор,
професор кафедри економіко-
математичного моделювання
та інформаційних технологій
Національний університет
«Острозька академія»

Октисюк Є.В.

студент
Національний університет
«Острозька академія»

Averkyna Maryna

The National University
of Ostroh Academy
Oktysiuk Yevhenii

The National University
of Ostroh Academy

Статтю присвячено моделюванню кредитоспроможності клієнта з використанням логістичної регресії. Сьогодні дуже багато українських банків і кредитних компаній зіштовхуються із ситуацією неповернення кредитів чи прострочок по виплатах. Така ситуація є негативним процесом, оскільки впливає на виконання однієї з найважливіших функцій банківської системи – акумулювання коштів та ефективний їх розподіл між суб'єктами економічної діяльності. Для цього у статті розглянуто підходи до визначення поняття «кредитоспроможність», проведено їх узагальнення, порівняння і класифікацію. Представлено авторське розуміння поняття «кредитоспроможність». Висвітлено особливості побудови логістичної регресії як методу оцінки кредитоспроможності клієнта; здійснено первинну роботу з даними і на їхній основі побудовано логістичну модель, яка дає можливість прогнозувати кредитоспроможність клієнта – фізичної особи банку. На основі результатів моделі зроблено висновки та надано відповідні рекомендації для комерційних банків.

Ключові слова: кредитоспроможність, регресія, логістична регресія, позичальник, дефолт, прострочка.

Стаття посвящена моделированию кредитоспособности клиента с использова-

нием логистической регрессии. Сегодня очень много украинских банков и кредитных компаний сталкиваются с ситуацией невозврата кредитов или просрочек по выплатам. Такая ситуация является негативным процессом, так как влияет на выполнение одной из важнейших функций банковской системы – аккумулярование средств и их эффективное распределение между субъектами экономической деятельности. Для этого в статье рассмотрены подходы к определению понятия кредитоспособности, проведено их обобщение, сравнение и классификация. Представлено авторское понимание концепции кредитоспособности. Освещены особенности построения логистической регрессии как метода оценки кредитоспособности клиента; осуществлена первичная работа с данными и на их основе построена логистическая модель, которая дает возможность прогнозировать кредитоспособность клиента – физического лица банка. На основании результатов модели сделаны выводы и даны соответствующие рекомендации для коммерческих банков.

Ключевые слова: кредитоспособность, регрессия, логистическая регрессия, заемщик, дефолт, просрочка.

A loan is an inseparable part of current economic system. The loaning is a main and one of the most dynamical activities, which banks do. Giving loans helps solve many social problems. For instance, improving life conditions, increasing payable demand and growing extent of gross domestic product. On the other hand as well banks can boost their financial results through gained incomes from money sold out. When a bank gives a loan it constantly tries to predict whether this client will pay back or default. So giving a loan is a risky business if a recipient doesn't want to pay back or can't. And if a bank does give a loan it loses money. Although there can be a situation when a bank doesn't give a loan to an honest client and finally it has alternative expenditure. Therefore banks need to boost predicting client's creditability. There have been many delinquencies and defaults in banking system of Ukraine recently. The banks face severe losses. Therefore it's vital to provide banks with reliable instrument for defining client's creditability in order economy function healthy as this instrument will boost financial results of the bank. On the other hand it will solve certain social problems. There many different methods to predict creditability. One of them is logistic regression. It is getting more and more popular recent time in Ukraine. Logistic regression is a sort of multiple regression that describes relationship between independent variables(factor variables) and dependent variables(resulting variable). Binary logistic regression is a regression with resulting variable that can have only two means. In our case it's default or non-default of a borrower. We got model with six variables in the process of modeling. We made the recommendations for a commercial bank based on each variable in the model. In the article there are considered approaches for defining the concept of creditability; theoretical aspects of logistic regression's building; the data was initially analyzed and built logistic model that predicts creditability of individual-client bank. There were made recommendations for a commercial bank based on model.

Key words: creditability, regression, logistic regression, borrower, default, delinquency.

Постановка проблеми. Ринковій економіці як економічній системі в Україні притаманна зрілість банківської системи і кредитних відносин. Розвиненість цієї ланки економічної системи уможливує вдосконалення і полегшення господарської діяльності для економічних суб'єктів шляхом надання грошей у користування. При цьому банківська установа, в яку приходять позичальник, стикається з дилемою – надати чи не надати кредит, оскільки, надавши кредит ненадійному позичальнику, який не зможе чи не захоче повернути кредит, банк утратить свої гроші, а не надавши

кредит хорошему позичальнику, банк утрачає можливості заробити гроші. Таким чином, перед банком постає проблема оцінки кредитоспроможності клієнта, тобто точна ідентифікація надійних і ненадійних клієнтів. Це зумовлює побудову відповідної моделі та визначення критеріїв до оцінювання кредитоспроможності клієнта.

Аналіз останніх досліджень і публікацій. Варто зазначити, що банківська система України порівняно з банківськими системами європейських країн чи Сполучених Штатів Америки є доволі молодою, тим не менше багато вітчизняних

науковців присвятили значну кількість своїх праць досліджуваній проблематиці. Дослідженням кредитоспроможності займалися такі вітчизняні вчені, як М. Белих, Л.Т. Гиляровська, Т.Д. Косова, О.І. Лаврушин, В.Д. Лагутін, Л.А. Лахтіонова, О.О. Олійник, А.І. Поезднік, Л.О. Примостка, В.Г. Федоренко, В.В. Хрестинін [4–7]. І.Г. Брітченко і О.Г. Момот, досліджуючи скоринг, визначили його як систему оцінки кредитоспроможності потенційного позичальника. Ця оцінка базується на числових статистичних методах, які застосовуються на різних етапах взаємовідносин між кредитором (банком) і позичальником (клієнтом). За результатами набраних балів скорингова система самостійно пропонує рішення про погодження чи відмову у видачі кредиту [1].

Г.І. Великоіваненко досліджувала наявні підходи до оцінювання кредитоспроможності позичальників комерційного банку, визначення адекватного математичного інструментарію моделювання скорингових систем щодо оцінювання кредитоспроможності позичальників комерційного банку [2].

А.М. Герасимович досліджує процес аналізу й оцінювання кредитоспроможності клієнта, а також аналізує якісні та кількісні показники, використовувані для оцінювання кредитоспроможності [3].

Таким чином, прослідковується глибоке теоретичне підґрунтя означеного питання. Проте вважаємо за необхідне констатувати малу кількість практичних напрацювань із виявлення реальних чинників, які впливають на кредитоспроможність позичальника – фізичної особи.

Постановка завдання. Метою статті є розкриття суті поняття «кредитоспроможність»; опис методу логістичної регресії, який використовується для моделювання кредитоспроможності; проведення моделювання і визначення основних чинників, які впливають на кредитоспроможність клієнта; надання конкретних рекомендацій для банків із метою поліпшення оцінки кредитоспроможності.

Виклад основного матеріалу дослідження. Під кредитоспроможністю ми розуміємо такий фінансовий стан позичальника, який дає змогу створювати грошові кошти і за їх рахунок повертати позички в повному обсязі й у визначений строк. На нашу думку, необхідно зосередити увагу на логістичній моделі бінарної залежної змінної. Це пояснюється тим, що в банках для визначення кредитоспроможності все більшої популярності набуває логістична регресія. Логістична регресія є різновидом множинної регресії, яка описує взаємозв'язок між незалежними змінними (факторними змінними) і залежною змінною (результуючою). Бінарна логістична регресія – це регресія, результуюча змінна якої може набувати лише двох значень. У нашому разі це настання чи ненастання дефолту позичальника.

Логістична модель має такий вигляд:

$$y = a_0 + \sum_{i=1}^n a_i x_i, \quad (1)$$

де $a_0 \dots a_1 \dots a_n$ – коефіцієнти при змінних, x_i – вхідні дані моделі, тобто дані про клієнтів, на основі яких банк робить висновок про їхню кредитоспроможність.

Для отримання значення ймовірності настання дефолту від 0 до 1 необхідно здійснити так зване логіт-перетворення, де:

$$P = \frac{1}{1 + e^{-y}} \quad (2)$$

P – імовірність настання події, яка цікавить дослідника;

e – основа натуральних логарифмів рівна 2,71...;

y – рівняння регресії.

Така операція дає змогу отримати вихідні значення ймовірності в проміжку [0; 1]. Проте для отримання даних про кредитоспроможність досліднику також необхідно власноруч увести порогові значення (cut-off) для розмежування між «поганими» і «хорошими» клієнтами [2].

Працівник банку повинен визначати такі типи помилок:

– TP (True Positives) – вірно класифіковані позитивні приклади;

– TN (True Negatives) – вірно класифіковані негативні приклади;

– FN (False Negatives) – позитивні приклади, класифіковані як негативні (помилка I роду);

– FP (False Positives) – негативні приклади, класифіковані як позитивні (помилка II роду) [9].

Далі модель аналізують за такими відносними показниками:

Чутливість (Sensitivity) – відношення кількості вірно класифікованих позитивних випадків до суми самих цих випадків і хибних негативних випадків:

$$S_e = \frac{TP}{TP + FN} \quad (3)$$

Специфічність (Specificity) – відношення кількості вірно класифікованих негативних ситуацій до суми цих випадків і хибних позитивних випадків:

$$S_p = \frac{TN}{TN + FP} \quad (4)$$

Частка позитивних ситуацій визначена правильно (Positive predictive value):

$$TPR = \frac{TP}{TP + FP} \quad (5)$$

Частка негативних ситуацій визначена правильно (Negative predictive value):

$$TNR = \frac{TN}{TN + FN} \quad (6)$$

Наостанок будується графік, який називається ROC-кривою (Receiver Operator Characteristic) і показує, наскільки модель якісна. Площу під кривою

називають AUC (Area Under Curve), чим її значення ближче до одиниці, тим точніше модель передбачає ймовірність настання дефолту [9]. Шкалу, яка дає змогу визначити якість моделі залежно від значень показника AUC, подано в табл. 1.

Таблиця 1
Експертна шкала значень AUC

Інтервал AUC	Якість моделі
0,9-1,0	Чудова
0,8-0,9	Дуже хороша
0,7-0,8	Хороша
0,6-0,7	Посередня
0,5-0,6	Незадовільна

Джерело: складено за [8]

У ході нашого дослідження для аналізу кредитоспроможності було використано дані про 2 661 клієнта одного з європейських банків. Дані знаходяться у вільному доступі на сайті Інтернет-спільноти дослідників даних Kaggle.

У руслі аналізу кредитоспроможності клієнта було вибрано такі характеристики:

- 1) обсяг позички;
- 2) сума застави, якою володіє позичальник;

- 3) кількість років, проведених на теперішній роботі;
- 4) кількість несхвальних звітів;
- 5) кількість прострочених кредитних ліній;
- 6) тривалість найдовшої кредитної лінії в місяцях;
- 7) кількість нещодавніх кредитних ліній.

Для того щоб визначити наявність статично значимих зв'язків між змінними, нами проведено кореляційний аналіз і отримано матрицю кореляційних даних (табл. 2). Як видно з даних матриці, статично значимий зв'язок наявний сумі застави, якою володіє позичальник (Сз), кількості років, проведених на теперішньому місці роботи (Р), з іншими факторними змінними.

Побудова логістичної регресії виконувалася в статистичному пакеті Stata/ IC 14.2. Показник Success Probability Cutoff, відповідно до якого класифікувалися дані, було встановлено на рівні 0,5. Як вхідні дані використано всі сім змінних, які представлено вище. Результати опрацювання первинних показників представлено на рис. 1.

Відповідно до даних рис. 3, показник кількості років на теперішній роботі має найнижче значення z, через що приймаємо рішення виключити його через несуттєвість впливу на залежну змінну.

```

Logistic regression                               Number of obs   =      2,661
LR chi2(7)                                       =      583.66
Prob > chi2                                       =      0.0000
Pseudo R2                                        =      0.2181
Log likelihood = -1046.366
    
```

Дефолтзалежна	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
сумапозички	-.0000967	.000014	-6.90	0.000	-.0001242 - .0000692
сумаіснуючоїзастави	-.0000106	2.24e-06	-4.75	0.000	-.000015 -6.24e-06
простроченікредитнілінії	.6285938	.0541632	11.61	0.000	.5224359 .7347517
вікнастарішоїлініївмісяцях	-.0068929	.000849	-8.12	0.000	-.0085569 -.0052289
суманещодавніхкредитнихліній	.1717167	.0320367	5.36	0.000	.1089259 .2345076
рокинатеперішнійроботі	-.0153438	.0082886	-1.85	0.064	-.0315892 .0009016
кількістьнесхвальнихзвітів	.6937769	.0778831	8.91	0.000	.5411288 .8464249
_cons	.8884177	.2474432	3.59	0.000	.403438 1.373397

Рис. 1. Коефіцієнти за змінних

Джерело: авторські розрахунки

Таблиця 2

Матриця кореляції вхідних даних

	Оп	Сз	Р	Кз	Пл	Тнл	Снл
Оп	1						
Сз	0,0888	1					
Р	-0,1087	-0,2971	1				
Кз	-0,0706	-0,1227	0,2284	1			
Пл	-0,0552	0,0499	0,3317	0,2315	1		
Тнл	-0,0115	0,0844	0,1461	-0,0296	0,0402	1	
Снл	-0,1337	-0,3586	0,0780	0,1828	0,1776	-0,1250	1

Джерело: авторські розрахунки

Нами було вирішено виключити ці змінні і провести тест із шістьма змінними, що залишилися, а результати представлено на рис. 2.

Відповідно до значення p-value, усі шість змінних попадають в інтервал значимості для $p=0,05$. Згідно із z-статистикою, змінні за впливом розташовуються у такій послідовності: кількість прострочених кредитних ліній, кількість несхвальних звітів, вік найстарішої лінії в місяцях, сума позички, сума нещодавніх кредитних ліній і сума існуючої застави.

Отримали класифікаційну матрицю (рис. 3), згідно з якою модель віднесла правильно 167 клієнтів, які повернули кредит, і також правильно класифікувала 2 054 клієнтів, які такі стикнулися з дефолтом.

Інші показники, які показують якість моделі дорівнюють таким значенням (рис. 4).

Нами визначено, що чутливість моделі становить 31,10%. Це низький показник, і таке його значення означає, що якщо модель має позитивний результат, вона є ненадійною. Специфічність

моделі становить 96,7%. Це значення близьке до ідеального й означає, що висновок про дефолт буде дуже надійним. Модель виявила, що частка клієнтів, які насправді повернуть кредит, становить 70,46%. Показник доволі непоганий, але його варто було б підсилити. Частка тих, хто точно виявлені як такі, що не повернуть кредит з усіх тих, хто зіштовхнеться з дефолтом, становить 84,74%. Цей показник є досить хорошим.

Таким чином, показник точності моделі становить 83,46%, що свідчить про модель як значущу для прийняття управлінських рішень. Модель передбачає дефолт позичальника краще, ніж його здатність повернути борг.

ROC-крива, яка показує, наскільки якісна модель, має такий вигляд (рис. 5).

Числове значення AUC, яке інтерпретує графік кривої, становить 0,8056, що, згідно з експертною шкалою значень, визначає модель як дуже хорошу. Отже, логістична модель кредитоспроможності клієнта матиме такий вигляд:

Logistic regression	Number of obs	=	2,661
	LR chi2(6)	=	580.18
	Prob > chi2	=	0.0000
Log likelihood = -1048.1086	Pseudo R2	=	0.2168

Дефолтзалежна	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
сумапозички	-.0000973	.000014	-6.94	0.000	-.0001247 -.0000698
сумаіснуючоїзастави	-.0000102	2.22e-06	-4.61	0.000	-.0000146 -5.89e-06
простроченікредитнілінії	.6191026	.0536709	11.54	0.000	.5139096 .7242955
вікнастарішоїлініївмісяцях	-.0072747	.0008262	-8.80	0.000	-.008894 -.0056553
суманещодавніхкредитнихліній	.1710377	.0317492	5.39	0.000	.1088103 .233265
кількістьнесхвальнихзвітів	.7060761	.0777426	9.08	0.000	.5537035 .8584487
_cons	.8084622	.2433129	3.32	0.001	.3315777 1.285347

Рис. 2. Значення коефіцієнтів за змінних

Джерел: авторські розрахунки

Logistic model for Дефолтзалежна

Classified	True		Total
	D	~D	
+	167	70	237
-	370	2054	2424
Total	537	2124	2661

Рис. 3. Класифікаційна матриця

Джерело: авторські розрахунки

Classified + if predicted Pr(D) >= .5
True D defined as Дефолтзалежна != 0

Sensitivity	Pr(+ D)	31.10%
Specificity	Pr(- ~D)	96.70%
Positive predictive value	Pr(D +)	70.46%
Negative predictive value	Pr(~D -)	84.74%
False + rate for true ~D	Pr(+ ~D)	3.30%
False - rate for true D	Pr(- D)	68.90%
False + rate for classified +	Pr(~D +)	29.54%
False - rate for classified -	Pr(D -)	15.26%
Correctly classified		83.46%

Рис. 4. Показники точності моделі

Джерело: авторські розрахунки

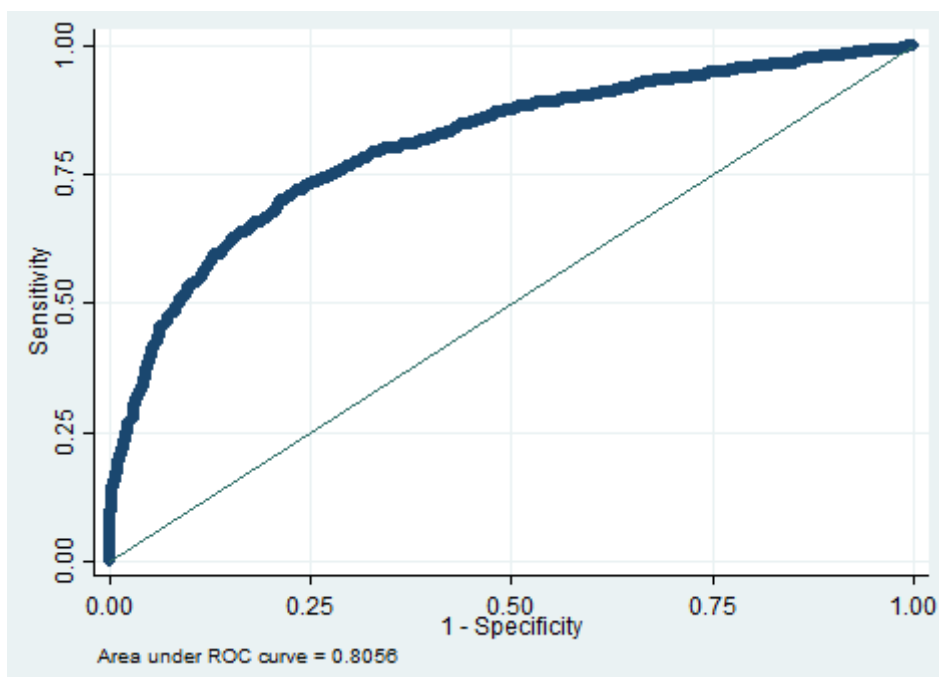


Рис. 5. ROC-крива

Джерело: авторські розрахунки

$$\begin{aligned}
 P(\text{дефолт}) = & 0,889 + 0,62 \cdot \text{пр. кр лін} + \\
 & + 0,71 \cdot \text{кл. несх. зв.} - 0,007 \cdot \text{найст. лінія} - \\
 & - 0,000097 \cdot \text{сума поз.} + 0,17 \cdot \text{нещ кред лінії} - \\
 & - 0,00001 \cdot \text{сума заст.}
 \end{aligned} \quad (7)$$

Висновки з проведеного дослідження. Отже, рекомендації для банку, на підґрунті яких формуватимуться управлінські рішення, здійснимо на основі логістичної регресії. У першу чергу банку слід звертати увагу на історію позичальника. Оскільки коефіцієнт за цією змінною є порівняно високим, то наявність хоча б однієї простроченої кредитної лінії надійно свідчить про ненадійність клієнта. Дуже малоймовірно, що такий клієнт поверне позичку.

Наступним за важливістю чинником є кількість несхвальних звітів про клієнта. По своїй суті він близький до попереднього показника, оскільки характеризує несхвальні звіти про клієнта в банку, пов'язані із запізнилими виплатами за позичкою. Якщо клієнт мав таку поведінку в минулому, то дуже малоймовірно, що він поліпшить свою поведінку в майбутньому. Найдовша лінія є третім за значимістю показником, на який банку слід звертати увагу. Знову ж таки тут маємо очікуване від'ємне значення при коефіцієнті. Чим тривалішу кредитну лінію клієнт відкривав у минулому, тим більш надійним він є. Сума позички має обернений вплив на прогнозування дефолту позичальника. Це означає, що чим більша сума позички, яку просить клієнт, тим менше банк сумнівається в його ненадійності. Пояснити це можна тим, що індивідууми психологічно не схильні просити великий кредит, якщо мають намір

не повертати його. Вони шукають такий баланс, щоб банк утратив таку суму, через яку не йому не прийдеться всерйоз займатися клієнтом, який не повернув кредит. П'ятим за значимістю є показник нещодавніх кредитних ліній, який має прямий вплив на прогнозування дефолту фізичної особи. Логічно, якщо клієнт має недавні позички, то навряд чи він устиг їх погасити і, можливо, шукає гроші, щоб погасити їх. Сума застави найменше впливає на результат прогнозування. Очікувано, що наявність застави свідчить на користь надійності клієнта. Отже, банку слід у першу чергу звертати увагу на показники, збільшення яких позитивно впливає на оцінку кредитоспроможності, а вже потім на показники з негативним впливом.

БІБЛІОГРАФІЧНИЙ СПИСОК:

1. Брітченко І.Г., Момот О.М. Особливості використання скорингових систем у банківському кредитуванні фізичних осіб. *Фінанси, облік і аудит*. 2011. № 17. С. 31–37.
2. Великоіваненко Г.І., Трокоз Л.О. Моделювання кредитоспроможності позичальників комерційного банку. *Наукові записки Національного університету «Острозька академія». Економіка*. 2013. № 22. С. 137–141. URL : http://nbuv.gov.ua/UJRN/Nznuoa_2013_22_31 (дата звернення: 26.12.2019).
3. Герасимович А.М. Проблеми аналітичної оцінки банківських методик визначення кредитоспроможності позичальника – фізичної особи. *Вісник Житомирського державного технологічного університету. Серія «Економічні науки»*. 2012. № 3(61). С. 313–315.

4. Гиляровская Л.Т. Экономический анализ : учебник для вузов / под. ред. Л.Т. Гиляровской. Москва : ЮНИТИ-ДАНА, 2002. 247 с.

5. Дудник А.Д. Словник – довідник менеджера по управлінню конкурентоспроможністю. Київ : КНЕУ, 1999. 60 с.

6. Зінченко О.А., Святенко С.В., Марчукова В.С. Узагальнення теоретичних підходів до визначення сутності категорії «кредитоспроможність». *Економіка. Управління. Інновації*. 2013. № 1. URL : http://nbuv.gov.ua/UJRN/eui_2013_1_23 (дата звернення: 26.12.2019).

7. Лахтіонова Л.А. Фінансовий аналіз сільськогосподарських підприємств : навчальний посібник. Київ : КНЕУ, 2003. 452 с.

8. Паклин Н. Логистическая регрессия и ROC анализ – математический аппарат. URL : <http://www.basegroup.ru/library/analysis/regression/logistic/> (дата звернення: 26.12.2019).

9. Davis J., Goadrich J. The Relationship Between PrecisionRecall and ROC Curves : Proc. Of 23 International Conference on Machine Learning, 2006/06/25 – 2006/06/29, Pittsburgh, PA, USA ACM Press.

REFERENCES:

1. Britchenko, I.H., Momot O. M. (2011) Osoblyvosti vykorystannja skorynghovykh system u bankivskomu kredytuvanni fizychnykh osib [Characteristics of using scoring systems in bank crediting of individuals]. *Finansy, oblik i audyt*, vol. 17, pp. 31-37.

2. Velykoivanenko H.I. (2013) Modeljuvannja kredytospromozhnosti pozycchalnykiv komercijnogho banku [Modelling creditability of a commercial bank's clients]. *Naukovi zapysky Natsional'noho universytetu "Ostroz'ka akademiiia"*. *Ekonomika* (electronic journal), vol. 22,

pp. 137-141. Available at: http://nbuv.gov.ua/UJRN/Nzn-uo_a_2013_22_31 (accessed 26 December 2019).

3. Herasymovych A.M. (2012) Problemy analitichnoji ocinky bankivskykh metodyk vyznachennja kredytospromozhnosti pozycchalnyka-fizychnoji osoby [Problems of analytical estimate bank methods for defining creditability individual borrower], *Visnyk Zhytomyr'skoho derzhavnogo tekhnologichnoho universytetu / Seriya: Ekonomichni nauky*, vol. 3(61), pp. 313–315.

4. Giljarovskaja L. T. (2002), *Jekonomicheskij analiz: uchebnik dlja vuzov* [Economic analysis: the manual for universities]. Moscow: JuNITI-DANA. (in Russian)

5. Dudnyk A. D. Slovyk (1999), *Dovidnyk menedzhera po upravlinniu konkurentospromozhnistiu* [The manager's guide to manage creditability]. Kyiv: KNEU. (in Ukrainian)

6. Zinchenko O. A. (2013) Uzagalnennja teoretichnykh pidkhodiv do vyznachennja sutnosti kategoriji «kredytospromozhnistj». [Generalization technological approaches to define the essence of the category «creditability»]. *Ekonomika. Upravlinnia. Innovatsii* (electronic journal), vol. 1. Available at: http://nbuv.gov.ua/UJRN/eui_2013_1_23 (accessed 26 December 2019).

7. Lakhtionova L. A. (2003), *Finansovyj analiz sil'skohospodars'kykh pidpriemstv: navch. posibnyk*, [Financial analysis of the agricultural enterprises: the manual], Kyiv: KNEU. (in Ukrainian)

8. Paklin N. (2006), Logisticheskaya regresiya i ROC analiz – matematicheskij aparat [Logistic regression i ROC-analysis – mathematical appliance], *Tehnologii analiza dannyh* (electronic journal), available at: <http://www.basegroup.ru/library/analysis/regression/logistic/> (accessed 26 December 2019)

9. Davis J., Goadrich J. (2016) *The Relationship Between PrecisionRecall and ROC Curves* : Proc. Of 23 International Conference on Machine Learning, 25.06 – 29.06, Pittsburgh, PA, USA ACM Press.