ҚАЗАҚСТАН РЕСПУБЛИКАСЫНЫҢ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РЕСПУБЛИКИ КАЗАХСТАН
MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE REPUBLIC OF KAZAKHSTAN

ХАЛЫҚАРАЛЫҚ УНИВЕРСИТЕТІ
МЕЖДУНАРОДНЫЙ УНИВЕРСИТЕТ
INTERNATIONAL UNIVERSITY

ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ

МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

2022 (3) 2
Сәуір-маусым

ISSN 2708–2032 (print)
ISSN 2708–2040 (online)

Халықаралық ақпараттық және коммуникациялық технологиялар журналы
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Меншіктенуші: «Халықаралық ақпараттық технологиялар университеті» АҚ (Алматы қ.).
Қазақстан Республикасы Ақпарат және әлеуметтік даму министрлігінің Ақпарат комитетінде – **20.02.2020** жылы берілген.
№ KZ82VPY00020475 мерзімдік басылым тіркеуіне қойылу туралы куәлік.
Тақырыптық бағыты: ақпараттық технологиялар, әлеуметтік-экономикалық жүйелерді дамытудағы цифрлық технологиялар, ақпараттық қауіпсіздік және коммуникациялық технологияларға арналған.
Мерзімділігі: жылына 4 рет.
Тиражы: 100 дана
Редакцияның мекенжайы: 050040, Алматы қ-сы, Манас к-сі, 34/1, 709-кабинет, тел: +7 (727) 244-51-09).
E-mail: ijict@iitu.edu.kz Журнал сайты: https://journal.iitu.edu.kz

Международный журнал информационных и коммуникационных технологий
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Собственник: АО «Международный университет информационных технологий» (г. Алматы).
Свидетельство о постановке на учет периодического печатного издания в Министерство информации и общественного развития Республики Казахстан **№ KZ82VPY00020475,** выданное от **20.02.2020 г.**
Тематическая направленность: информационные технологии, информационная безопасность и коммуникационные технологии, цифровые технологии в развитии социо-экономических систем**.**
Периодичность: 4 раза в год.
Тираж: 100 экземпляров.
Адрес редакции: 050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09).
E-mail: ijict@iitu.edu.kz Сайт журнала: https://journal.iitu.edu.kz

---

«International Journal of Information and Communication Technologies»
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Owner: International Information Technology University JSC (Almaty).
The certificate of registration of a periodical printed publication in the Ministry of Information and Social Development of the Republic of Kazakhstan, Information Committee No. KZ82VPY00020475, issued on 20.02.2020.
Thematic focus: information technology, digital technologies in the development of socio-economic systems, information security and communication technologies
Periodicity: 4 times a year.
Circulation: 100 copies.
Editorial address: 050040. Manas st. 34/1, Almaty. +7 (727) 244-51-09). E-mail: ijict@iitu.edu.kz
Journal website: https://journal.iitu.edu.kz

МАЗМҰНЫ

СОДЕРЖАНИЕ

CONTENTS

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 3. Is. 2. Number 10 (2022). Pp. 103–113
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/*10.54309/IJICT.2022.10.2.010*

UDC 519.257, 005

RESEARCH ON RISK ANALYSIS METHODS USING MODELS OF DEFAULT PROBABILITY IN THE FINANCIAL INDUSTRY

A.R. Suleimenova[1*], A.Zh. Sayabayeva[1], A.N. Moldagulova[2]
[1]International Information Technology University, Almaty, Kazakhstan;
[2]Kazakh National Research Technical University, Almaty, Kazakhstan.

Modagulova Aiman Nikolayevna — Candidate of Physical and Mathematical Sciences, Associate Professor, Head of the Department "Software Engineering", Kazakh National Research Technical University named after K.I. Satpayev
ORCID:0000-0002-1596-561X. E-mail: aigerimsuleimenova98@gmail.com;
Suleimenova Aigerim Raikhanovna — Master's student, specialty "Business Analytics", the International University of Information Technologies
ORCID: 0000-0002-5754-651X;
Sayabayeva Arida Zhomartkyzy — Master's student, specialty "Business Analytics", the International University of Information Technologies
ORCID:0000-0003-4403-7625. E-mail: arida0304z@gmail.com.

Abstract. The article discusses the issues of methods for analyzing the probability of default (bankruptcy) of enterprises in the context of the financial industry. Based on the research, it proposes approaches to assess the probability of default. Many models help analyze credit risks, for example, the probability of default, migration risk, and default on losses. Each of these models is vital for assessing credit risk, however, one of the basic and essential models is the probability of the default model, i.e., used in this article. To solve this problem, a standard mathematical apparatus is used: linear regression, matrix theory, and nonlinear programming.

Keywords: the probability of default; loan quality category; reserves for possible loan losses; binary regression

For citation: A.R. Suleimenova, A.Zh. Sayabayeva, A.N. Moldagulova. Research on risk analysis methods using models of default probability in the financial industry // INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES. 2022. Vol. 3. Is. 2. Number 10. Pp. 103–113 (In Russ.). DOI: *10.54309/IJICT.2022.10.2.010.*

ҚАРЖЫ САЛАСЫНДАҒЫ ТӘУЕКЕЛДЕРДІ ТАЛДАУ ӘДІСТЕРІН ҚАРЖЫ САЛАСЫНДАҒЫ ҚАРЖЫЛЫҚ ЫҚТИМАЛДЫҚ ҮЛГІЛЕРІН ПАЙДАЛАНҒАН ЗЕРТТЕУ

А.Н. Молдагулова[1]*, А.Р. Сулейменова[1], А.Ж. Саябаева[2]

Молдагулова Айман Николаевна — физика-математика ғылымдарының кандидаты, доцент, Қ.И. Сәтбаев атындағы Қазақ ұлттық ғылыми-зерттеу техникалық университеті «Программалық инженерия» кафедрасының меңгерушісі
ORCID:0000-0002-1596-561X. E-mail: aigerimsuleimenova98@gmail.com;
Сулейменова Айгерим Райхановна — Халықаралық ақпараттық технологиялар университетінің «Бизнес-аналитика» мамандығының екінші курс магистранты
ORCID: 0000-0002-5754-651X.
Саябаева Арида Жомартқызы — Халықаралық ақпараттық технологиялар университетінің «Бизнес-аналитика» мамандығының екінші курс магистранты
ORCID:0000-0003-4403-7625. E-mail: arida0304z@gmail.com.

Аннотация. Мақалада қаржылық индустрия контекстінде кәсіпорындардың дефолт (банкрот) ықтималдығын талдау әдістері қарастырылады. Зерттеу негізінде ол дефолт ықтималдығын бағалау тәсілдерін ұсынады. Көптеген модельдер несиелік тәуекелдерді талдауға көмектеседі, мысалы, дефолт ықтималдығы, көші-қон тәуекелі және шығындар бойынша дефолт. Осы үлгілердің әрқайсысы несиелік тәуекелді бағалау үшін өте маңызды, дегенмен негізгі және маңызды үлгілердің бірі әдепкі үлгінің ықтималдығы болып табылады, яғни осы мақалада пайдаланылады. Бұл мәселені шешу үшін стандартты математикалық аппарат қолданылады: сызықтық регрессия, матрицалық теория және сызықты емес бағдарламалау.

Түйін сөздер: дефолт ықтималдығы, несие сапасы санаты, несие бойынша ықтимал шығындарға арналған резервтер, екілік регрессия

Дәйексөз үшін: Молдагулова А.Н., Сулейменова А.Р., Саябаева А.Ж. Қаржы саласындағы тәуекелдерді талдау әдістерін қаржы саласындағы қаржылық ықтималдық үлгілерін пайдаланған зерттеу // ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ. 2022. Том. 3. Is. 2. Нөмірі 10. 103–113 бет (орыс тілінде). DOI: 10.54309/IJICT.2022.10.2.010.

ИССЛЕДОВАНИЕ МЕТОДОВ АНАЛИЗА РИСКА С ИСПОЛЬЗОВАНИЕМ МОДЕЛЕЙ ВЕРОЯТНОСТИ ДЕФОЛТА В ФИНАНСОВОЙ ОТРАСЛИ

А.Н. Молдагулова[1]*, А.Р. Сулейменова[1], А.Ж. Саябаева[2]

Молдагулова Айман Николаевна — кандидат физико-математических наук, ассоциированный профессор, заведующая кафедрой «Программная инженерия», Казахского национального исследовательского технического университета имени К.И. Сатпаева

ORCID:0000-0002-1596-561X. E-mail: aigerimsuleimenova98@gmail.com;
Сулейменова Айгерим Райхановна — магистрант второго курса специальности «Бизнес аналитика» Международного университета информационных технологий
ORCID:0000-0002-5754-651X;
Саябаева Арида Жомартқызы — магистрант второго курса специальности «Бизнес аналитика» Международного университета информационных технологий
ORCID:0000-0003-4403-7625. E-mail: arida0304z@gmail.com.

Аннотация. В статье рассматриваются вопросы методов анализа вероятности дефолта (банкротства) предприятий в разрезе финансовой отрасли. На основе исследования предлагаются подходы к оценке вероятности дефолта. Многие модели помогают анализировать кредитные риски, например вероятность дефолта, миграционный риск и дефолт по убыткам. Каждая из этих моделей жизненно необходима для оценки кредитного риска, однако одной из основных и существенных моделей является модель вероятности дефолта, т.е. используемая в данной статье. Для решения этой задачи используется стандартный математический аппарат: линейная регрессия, теория матриц и нелинейное программирование.

Ключевые слова: вероятность дефолта, категория качества кредита, резервы на возможные потери по ссудам, бинарная регрессия

Для цитирования: Молдагулова А.Н., Сулейменова А.Р., Саябаева А.Ж. Исследование методов анализа риска с использованием моделей вероятности дефолта в финансовой отрасли // МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ. 2022. Том. 3. Is. 2. Номер 10. Стр. 103–113 (на русском языке). DOI: 10.54309/IJICT.2022.10.2.010.

Introduction

To date, commercial banks in the developed countries have designed and tested many mathematical models for assessing borrowers' credit risk. Many of them have stood the test of time and are now used in world practice. Thus, the Basel Committee on Banking Supervision proposes a standardized approach for assessing credit risk, based on the assessment of external rating agencies, and an approach based on internal ratings (Internal Rating Based approach, IRB) (http://www.bis.org/publ/bcbsc111.htm).

The problem of solving issues related to default in the financial sphere is considered by us based on constructing a mathematical integral model that evaluates the individual risk of the borrower and acts as a basis for assessing the aggregate risk. The main disadvantage of this approach is the unaccounted fact that over the life of the loan there is a change in the individual probability of going into default. The actual assessment of the loan is investigated not only for the first year of the loan's life but also for the entire term of its life. According to the IFRS standard, the volume of reserves formed should depend on the dynamics of the risk level, which is determined by assessing the risk of default throughout the life of the loan (lifetime PD). A model is proposed for obtaining a dynamic estimate of the probability of default based on its data on the quality of

fulfillment of obligations under the loan agreement (future engineering) and the phase of the macroeconomic cycle. The model is based on the method of logistic regression with regularization, the quality of which is determined through the calculation of the Gini coefficient. Building a model for estimating the probability of going into default.

Methods

An essential class is market models that are based on information available to stock market players. Firstly, it is the market data of the borrower's listed securities. Such models can be divided into structural and abbreviated models.

At the center of the structure is the idea that the value of the company's shares is an option on the company's assets with a transaction price equal to its liabilities. The founders of this class of models are BLACK and Shoals, Merton. Merton considered the company's accounts payable as a claim that could be changed in its value and used the BLACK — S choles option pricing formula to assess the probability of the company's default (Black et al., 1973: 637–654). Within the framework of this model, the provision of a loan is interpreted as the purchase of company assets from shareholders and the transfer of an option on these assets with an exercise price equal to the cost of the loan and an execution time equal to the repayment of the loan. VT represents the value of the company's assets at time t, and after the payment expires at time T, creditors demand the nominal amount of debt D. the model uses a simplified description of the debt structure: it is assumed that the company has one debt in the amount of D in the form of a zero-coupon bond, that is, a zero-coupon debt. the total amount of debt must be paid in a lump sum at the time of T. Suppose that the amount of debt is fully secured by the assets of the company, that is, $VT \geq D$, then the shareholders receive the difference VT-D. If the debt is not secured by the assets of the enterprise, that is, the creditors receive what is available to them, the shareholders do not receive anything (this follows from the priority of fulfilling creditors' claims to shareholders in the event of bankruptcy of the enterprise).

With increasing uncertainty in financial markets, credit risk management is a priority in banking risk management. By the Basel Agreement, the basis of credit risk management is the formation of a reserve capitalization corresponding to the expected losses calculated for the portfolio (Karminsky et al., 2005). Therefore, the correctness of the assessment of expected losses becomes more relevant.

Models that are based on financial and accounting data of corporate borrowers of the bank are quite common. Depending on the statistical method used, they can be divided into scoring models, linear discriminant analysis models (among which, in turn, discriminant analysis models with one and several variables differ), and binary choice models. A credit score is a statistical method proposed by the American scientist David Durand in 1941, which was originally used to rank retail borrowers. When using the scoring model, each borrower is assigned a rating that characterizes his financial condition and ability to repay his obligations to the lender promptly. In the future, the entire range of possible values is divided into intervals, rating groups. By calibrating the model based on historical data, each rating score is adjusted to take into account the probability of default, which for the most part is a determination of

the proportion of companies in this group that have defaulted during the year. Now, credit-scoring models are quite common in banking practice, especially for assessing the probability of default on single retail loan portfolios. However, Durand's original scoring model was quite simplistic. The author took into account certain characteristics of each retail borrower (such as gender, age, living in the place, profession, place and work experience, financial condition), and depending on their cost, he assigned each client a certain point. If the total score for all characteristics exceeded 1.25 points, the client was considered solvent. Modern credit rating systems of Moderna have become much more complex and universal, but their disadvantage is discretion, as well as a prerequisite for having a wide base of credit histories. Since valuation models need a lot of background information about loans and borrowers' credit histories to determine dependencies, it is necessary to regularly update the reference data and, consequently, the identified dependencies. In the group of discriminant analysis models, discriminant analysis models with one variable, the ancestor of which is Beaver, are the simplest for assessing the borrower's default.

According to Basel II and IFRS, the assessment of the expected losses of the loan portfolio is based on the IRB approach to credit risk, which, if the bank complies with certain minimum conditions and disclosure requirements, allows it to rely on its internal assessments of risk components when calculating capital coverage of a certain risk [4]. According to the Basel II agreement, the calculation of expected losses EL (Expected Loss) on the loan is made according to the following formula:

$$EL = PD \cdot EAD \cdot LGD, \quad (1)$$

Where *PD* (Probability of Default) — the probability of default.

EAD (Exposure at Default) — exposure to credit risk, which is an economic assessment of the value of assets at risk at the time of default exposure to credit risk, which is an economic assessment of the value of assets at risk at the time of default;

LGD (Loss Given Default) — losses in case of default, reflecting the share of irretrievable loss in case of default. losses in case of default, reflecting the share of irretrievable loss in case of default.

At the same time, most often, like PD, an individual assessment of the probability of default is used, which is calculated based on a scoring model. According to the recommendations of the Basel II Agreement, overdue principal debt or interest for 90 days or more during the first year of the loan term is considered unpaid, and the probability of overdue principal debt or interest for more than 90 days during the first year of the loan term is defined as overdue (in the future, the the year will be equal to 365 calendar days). The main problem when using the current definition is the unregistered fact that over the life of the loan there is a change in the individual probability of default. In this regard, the issue of credit assessment not only for the first year of the loan term but also for its entire life is becoming more and more relevant. This provision is enshrined in the standard IFRS 9, according to which the number of reserves created should depend on the dynamics of the risk level, which is determined by the assessment of the risk of default on the loan during the entire loan term (lifetime estimation).

Loan portfolio data for the last month for a certain period (for example, 2012–01–

31, 2012–02–29, 2012–03–31, ...). For each cutting date t, a loan portfolio of assets is formed: loans opened for which there was no violation event and for which the actual values above the output violation indicator in the interval are known. Loan portfolios for each cut-off point t in one sample. The division into a training sample and a validation sample is based on a random selection of credits relative to 70/30 so that each credit falls only in the training sample or validation sample. It is worth noting that when using a simple random selection of observations, we will get an obvious retrain of the model: the coefficient in the test sample may be greater than in the training sample. The correct construction of training and validation samples will allow obtaining more stable estimates of models, which after the implementation of the model will exclude the occurrence of unpredictable situations when assessing the probability of default of the loan portfolio in any segment (provided that the portfolio structure will not change, and the macroeconomic situation will be stable).

The variables that will be used to build a behavior assessment model are based on the credit history of each loan. The raised issue of choosing variables that will correctly reflect the quality of fulfillment of obligations under the loan agreement (future engineering) is relevant. Let's introduce the following notation: let PR be a set of values for the number of days of delay that the client allows, and M is a set of values for the number of months during which the presence of a delay is estimated.

Then it is proposed to use the following quantitative estimates as variables:

- the number of exits for overdue PR in the first m months of the loan term -these variables allow you to assess the quality of loan servicing in the first year of the loan term;
- the number of exits on overdue PR in the last m months of the loan term (the countdown of months begins with the court date, timer) — allows you to assess the quality of credit service in the last year of the loan life before the court date, thus, the credit history of the last loan is evaluated;
- the ratio of the total number of outflows of overdue public relations for the entire loan term to the loan term in months;
- the ratio of the total number of exits from PR delinquencies for the entire term of the loan to the number of months that the loan has lived on the cut-off date;
- the proportion of the maximum number of months during which the loan was in permanent public relations debt, from the number of months during which the loan lived until the cut-off date.

Similarly, variables are calculated based on the overdue amount of loans:

- the ratio of the overdue PR amount in the first m months of the loan's life to the total loan amount;
- the ratio of the overdue amount of PR for the last m months of the life of the loan, to the total amount of the loan, the months are counted from the countdown cut-off date;
- the ratio between the total amount of overdue public relations debt during the entire term of the loan, divided into months, and the total amount of the loan;
- the ratio of the total amount of overdue PR for the entire term of the loan, divided

by the number of months during which the loan exists at the cut-off date, to the total amount of the loan;

- the ratio between the average amount overdue when the loan was in permanent public relations debt and the total amount of the loan.

The described variables can be attributed to a block of variables that change over the life of the loan. In addition, it is worth taking into account the fact that the historical sample of the loan portfolio from some segment T contains loans with different useful lives. Using the credit history of loans as variables, we get that there is incomplete and time-limited data for building the model, i.e. censored data on the right, i.e. there is some history of the customer's behavior on the loan before the cutoff, but there is no information about how the customer's behavior will look in the next period [t, T], where T is the cutoff time, T is the loan term. Because of this, an additional variable is introduced: the fact that the loan is valid until a certain point in time.

In addition to the block of variables that change over the life of the loan, it is possible to allocate a block of variables that are fixed at the time of loan issuance, namely:

- part of the initial payment of the amount of the loan issued;
- advance payment of the cost of the goods;
- PD0 score, which determines the probability of default during the first year of the loan's life and is calculated at the time of loan issuance based on the characteristics of the borrower's questionnaire. At the same time, most often, like PD, an individual assessment of the probability of default is used, which is calculated based on a scoring model. By the recommendations of the Basel II Agreement, overdue principal debt or interest for 90 days or more during the first year of the loan term is considered unpaid, and the probability of overdue principal debt or interest for more than 90 days during the first year of the loan term is defined as overdue (in the future, the year will be equal to 365 calendar days). The main problem when using the current definition is the unregistered fact that over the life of the loan there is a change in the individual probability of default. In this regard, the issue of credit assessment not only for the first year of the loan term but also for its entire life is becoming more and more relevant. This provision is enshrined in the standard IFRS 9, according to which the number of reserves created should depend on the dynamics of the risk level, which is determined by the assessment of the risk of default on the loan during the entire loan term (lifetime estimation).

Loan portfolio data for the last month for a certain period (for example, 2012–01–31, 2012–02–29, 2012–03–31, ...). For each cutting date t, a loan portfolio of assets is formed: loans opened for which there was no violation event and for which the actual values above the output violation indicator in the interval are known. Loan portfolios for each cut-off point t in one sample. The division into a training sample and a validation sample is based on a random selection of credits relative to 70/30 so that each credit falls only in the training sample or validation sample. It is worth noting that when using a simple random selection of observations, we will get an obvious retrain of the model: the coefficient in the test sample may be greater than in the training sample (Altman,

2003: 201–214). The correct construction of training and validation samples will allow obtaining more stable estimates of models, which after the implementation of the model will exclude the occurrence of unpredictable situations when assessing the probability of default of the loan portfolio in any segment (provided that the portfolio structure will not change and the macroeconomic situation will be stable).

Result

The variables that will be used to build a behavior assessment model are based on the credit history of each loan. The raised issue of choosing variables that will correctly reflect the quality of fulfillment of obligations under the loan agreement (future engineering) is relevant. Let us introduce the following notation: let PR be a set of values for the number of days of delay that the client allows, and M is a set of values for the number of months during which the presence of a delay is estimated.

Then it is proposed to use the following quantitative estimates as variables:

- the number of exits for overdue PR in the first m months of the loan term -these variables allow you to assess the quality of loan servicing in the first year of the loan term;
- the number of exits on overdue PR in the last m months of the loan term (the countdown of months begins with the court date, timer) - allows you to assess the quality of credit service in the last year of the loan life before the court date, thus, the credit history of the last loan is evaluated;
- the ratio of the total number of outflows of overdue public relations for the entire loan term to the loan term in months;
- the ratio of the total number of exits from PR delinquencies for the entire term of the loan to the number of months that the loan has lived on the cut-off date;
- the proportion of the maximum number of months during which the loan was in permanent public relations debt, from the number of months during which the loan lived until the cut-off date.

Similarly, variables are calculated based on the overdue amount of loans:

- the ratio of the overdue PR amount in the first m months of the loan's life to the total loan amount;
- the ratio of the overdue amount of PR for the last m months of the life of the loan, to the total amount of the loan, the months are counted from the countdown cut-off date;
- the ratio between the total amount of overdue public relations debt during the entire term of the loan, divided into months, and the total amount of the loan;
- the ratio of the total amount of overdue PR for the entire term of the loan, divided by the number of months during which the loan exists at the cut-off date, to the total amount of the loan;
- the ratio between the average amount overdue when the loan was in permanent public relations debt and the total amount of the loan.

The described variables can be attributed to a block of variables that change over the life of the loan. In addition, it is worth taking into account the fact that the historical sample of the loan portfolio from some segment T contains loans with different useful

lives. Using the credit history of loans as variables, we get that there is incomplete and time-limited data for building the model, i.e. censored data on the right, i.e. there is some history of the customer's behavior on the loan before the cutoff, but there is no information about how the customer's behavior will look in the next period [t, T], where T is the cutoff time, T is the loan term. Because of this, an additional variable is introduced: the fact that the loan is valid until a certain point in time.

In addition to the block of variables that change over the life of the loan, it is possible to allocate a block of variables that are fixed at the time of loan issuance, namely:

- part of the initial payment of the amount of the loan issued;
- advance payment of the cost of the goods;
- PD0 score, which determines the probability of default during the first year of the loan's life and is calculated at the time of loan issuance based on the characteristics of the borrower's questionnaire.

Analysis

Thus, variables are calculated for each contract, based on which a behavioral scoring model for assessing the default is built. From the available variety of model construction methods in the current example, the method of logistic regression with regularization was used as a method of model construction. The formula for calculating the scoring score PD:

$$PD = \frac{w}{1 + e^{-z}} \quad (2)$$

where z is calculated as the sum of the constant and the points corresponding to the characteristics of the borrower and their combinations:

$z = b_1 \cdot x_1 + b_2 \cdot x_2 + \ldots + b_n \cdot x_n + b_0$,

$x_1, x_2, \ldots, x_n$ –, values of variables

$b_1, b_2, \ldots, b_n$ – coefficients for variables,

b_0 –, some constant

w – correction factor for the macroeconomic cycle.

The z value is determined by the calculated variables, which are the characteristics of the borrower. Additional constraints are applied to the vector of weights in the model by regularizing the model. This approach consists in choosing small weights in the absolute value on average, which leads to less instability of the model, that is, dependence on training data. The elastic network method is used as a regularization scheme, which combines loop and ridge, regression models (Altman, 1968).

Discussion

Let us consider the construction of a model for estimating the probability of PDT default at the moment (slice) t on the example of a regional retail bank. The data of the loan portfolio for the last month, taking into account the fact of the "survival" of the loan to each segment understudy, serve as reference data for constructing a model for estimating the PDT score. The model is built in the recent segment of the portfolio. For

each cut-off date, a portfolio of current loans was formed-loans opened from the cut-off date and with a current default on principal or interest of no more than 15 days. The portfolios obtained for each cut-off point were combined into one training sample by the described method. The volume of the study sample amounted to 151586 loans, the share of overdue loans in which 0.04146 by 31–10–2017. The volume of the validation sample is 66310, the share of overdue loans is 0.04340. The sample size is 22038, the share of overdue loans is 0.03825. The division of the sample into formation and validation is based on the rule described above. For each of the samples, the above variables were calculated, assuming PR 1, 15, 30, 90, and M {1, 3, 6, 9, 12}, and based on it, a logistic regression model with regularization was built. The coefficient in the training sample was 0.7163, in the validation sample - 0.7024, in the test sample - the actual score according to the state (Figure 1). The z value is determined by the calculated variables, which are the characteristics of the borrower. Additional constraints are applied to the vector of weights in the model by regularizing the model. This approach consists in choosing small weights in the absolute value on average, which leads to less instability of the model, that is, dependence on training data. The elastic network method is used as a regularization scheme, which combines loop and ridge, regression models.

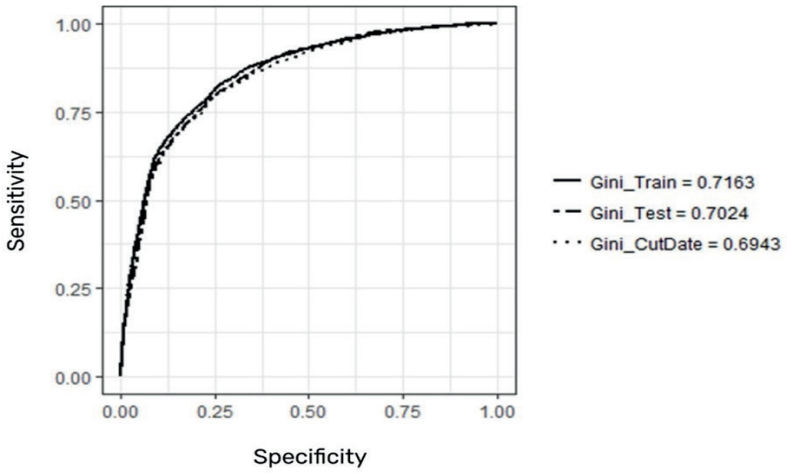

Figure 1 - *ROC curve graph*

Conclusion

In this paper, a solution to the problem was presented not only for the first year of the loan "life", but also for the entire lifetime (life, assessment). Based on the behavioral model of architecture evaluation. For instance, the regional retail bank used a car as an example of how to build a PDt default probability model to estimate at some point (slice) t, taking into account the terms of the loan agreement of fulfillment of obligations, the quality estimated by the described variable sizes (future engineering) and the macroeconomic phase of the cycle. The obtained quality indicators from the model indicate its high-power prediction. This credit risk assessment of the credit risk

during the action complies with the requirements of the standard and allows to create a reserve volume of the portfolio based on the level of risk in dynamics.

REFERENCES

Altman E.I. (2003). «Managing credit risk: a challenge for the new millennium». Economic Notes. — Vol. 31. — Issue 2 (December), — Pp. 201–214.

Altman E.I. (1968). «Financial rations. Discriminant analysis, and the prediction of corporate bankruptcy». Journal of Finance, September.

Basel Committee on Banking Supervision, International Convergence of Capital Measurement and Capital Standards, Comprehensive Version. — http://www.bis.org/publ/bcbsc111.htm.

Black F., Scholes M. (1973). «The pricing of options and corporate liabilities». The Journal of Political Economy. —Vol. 81. — No. 3 (May — Jun.). — Pp. 637–654.

Karminsky A.M., Peresetsky A.A., Petrov A.E. (2005). Ratings in economics: methodology and practice: Monograph / Edited by A.M. Karminsky. - M.: Finance and Statistics, — 2005.

Bobyshev A., Galperin F., Mishchenko Ya. (2005). The practice of applying the VaR methodology for assessing and managing credit risk in Alfa-Bank // Financial Risk Management. — 2005. — № 2.

ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ

МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

Правила оформления статьи для публикации в журнале на сайте:
https://journal.iitu.edu.kz
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Собственник: АО «Международный университет информационных технологий» (Казахстан, Алматы)

ОТВЕТСТВЕННЫЙ РЕДАКТОР
Ералы Диана Русланқызы

КОМПЬЮТЕРНАЯ ВЕРСТКА
Жадыранова Гульнур Даутбековна

Подписано в печать 15.06.2022.
Формат 60x881/8. Бумага офсетная. Печать - ризограф.7,0 п.л. Тираж 100
050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09.

Издание Международного университета информационных технологий
Издательский центр КБТУ, Алматы, ул. Толе би, 59