УДК 005

Bulantayev A.M. [1], Musakhan K.B. [1], Moldagulova A.N. [1], Sembina G.K. [1]

[1] International Information Technology University, Almaty, Kazakhstan

FORECASTING EXPECTED BANK LOSSES AT GRANTING A LOAN

Abstract. *This article uses the sample data of the SAS platform as an example to introduce the statistical analysis and prediction of the expected loss of loans issued by banks. The original data for this study comes from a Kaggle source, which provides information about the credit history of bank customers. The technology is based on logistic regression, graphical data analysis, and the basis of building a model on the SAS platform. The model can be used to predict credit risk and describe credit risk in the banking system.*

Keywords: *component, data analysis, credit risk, Loss Given Default, Expected Loss, Probability of Default, Exposure at Default, logistic regression, model, non-performing loans, Special Air Service platform, forecasting.*

Introduction

This kind of credit risk is one of the most common types of risks, and it has a vital impact on the robustness of a single bank's exercise and the entire framework of the currency of custody. One of the main reasons for emergencies is to prompt financial analysts to underestimate the risks associated with the use of unused currency violations and the subsequent emergencies. In combination with the potentially unsafe consequences of credit opportunities, it is essential to fully evaluate them.

Under this circumstance, the state-of-the-art bank must not determine the importance of its image in the same way that it assesses the opportunity level of a credit asset portfolio when conducting lending activities. In the unstable budget market and currency-related emergencies, it is vital to be able to foresee the pattern of bank marks.

However, the current problem is the need for mandatory tools to predict credit risk. We studied a modern method based on the use of evaluable strategies to estimate the expected misfortunes of banks when issuing advances, taking into account different variables that affect the risk of advances portfolio, and using information provided by Kaggle.

The biggest purpose of consideration is to estimate the unfortunate circumstances of banks when issuing credit for explanatory considerations of suitability. The strange thing to consider is the improvement of the program using the fact-finding data of the banking framework obtained by Kaggle. The emergence of term papers is a demonstration of the strategy of using historical bank data and the use of SAS equipment to analyze the credit risk within the scope of banking services.

New impetus for adjusting credit risk management

From the perspective of cash-related instructions, the conditions for the activation of the COVID-19 emergency have special propositions for monitoring and coordinating credit risk. In the past three months, banks have been adapting to the evolving trend and researching the most cutting-edge methods to deal with challenges. These inspections assess the overall impact of the emergency on the national or regional economy, and the impact of various departments and sub-sectors, especially credit risk issues that require real-time inspections. The examination started when it was directly discovered that this emergency situation had five curious effects on credit risk.

At the time of crisis, the change in budget conservatism is more critical than the differences between departments and the past divestments. Certain companies, such as nutrition wholesalers, dominate the crisis and are fighting to meet the increasing demand. Other factors, such as broadcast communications and medicines, were hardly affected. But as we all know, certain sectors (such as travel, transportation, tourism and neighborhoods) are indeed challenged [1].

Nonetheless, to properly outline cash-related robustness within this emergency, banks must go through inspections of parts or sub-sectors and investigate the borrowers. The business model from one company to another in the same sub-industry may be absolutely correct, and along these lines of thinking, it is more or less suitable for survival and faster recovery in the current environment. For example, many companies have very close online businesses, while others do not. Therefore, banks cannot draw conclusions from sub-industry inspections alone, that is, whether a particular borrower is annoyed.

In addition, the standard data sources used in credit risk assessment must become obsolete overnight. This crisis has manifested itself as a powerful external shock in most cases of the global credit cycle. Both supply and asking prices were accidentally stifled. In addition, suddenly, the data of the past six or twelve months of data is not important for the quality of the survey of individual borrowers. The innovative method of acquiring and using high-frequency data is today's organization [2].

In the process of developing the COVID-19 crisis, due to the prejudice of customers, their banks eventually developed naturally, and the benefits of habitual installment payment collection technologies (phones, emails, letters) are getting smaller and smaller. The huge wave of undesirable progress that is currently forming will occupy a long exhaust direction resource in the near future.

Towards Data-Driven analysis and Real-Time decision-making

In order to respond to emergency situations, teaching related to driving money has begun to adopt unused market segment surveys, borrower versatility and high-frequency analysis methods to recognize and observe. One of the key changes we have seen is that pioneers are rapidly shifting from the perspective of one department to a sub-sector, and finally using real-time information and analysis techniques of debt to quickly transfer to the perspective of debt [3]. Most banks have conducted precise speculations in specific sub-sectors and are getting closer (or up to now) to the debtor's view of risk assessment.

Crossing this skyline is the method of investigating forms of credit underwriting by using real-time trade information in decision-making and schedule analysis [4]. Switching to these unused strategies will help banks adapt to emergencies, but in addition, as a practice, we believe that credit risk management should be changed in the coming months and a long time. The best banks will indeed continue to maintain and extend these conditions after emergencies, in order to more effectively monitor credit risk, to serve senior customers and enable them to recover faster.

SAS application style for predicting expected bank losses when issuing loans

The software market continues to grow. A new software package was developed based on the rapidly developing modern computer technology. The functions of the program are being expanded so that people can use them not only as simple tools to simplify work, but also as mature assistants that can solve complex problems in banks. This study shows the use of statistical software package to analyze and predict bank data.

In order to determine the initial predicted value of the credit risk level, we recommend using SAS statistical analysis methods. To predict the risk of the loan portfolio, it is necessary to determine the criteria for changing the level of this indicator. We suggest that the main criterion for credit risk indicators is a model built on the bank's historical database, which shows the possible risks of a specific customer, that is, the possible loss of the bank due to a specific customer. We believe that this model most accurately characterizes the quality of bank loan portfolios.

All customer credit records are stored in the bank database. Credit history is information about your loan obligations. It shows the bank, the microfinance organization (MFI) or the credit consumer cooperative (CCP) from which you are applying for loans and borrowing. When was that and how much you spent. Have you ever been a co-borrower or guarantor of loans from others [5]? Do they pay carefully or delay payment?

The amount of loss that the bank may face can be predicted based on the credit history data. The average level of bank losses is estimated by calculating the expected loss index (EL). From the perspective of the bank, the indicator of expected loss is an element of business value. Use basic risk parameters to estimate the amount of expected loss: PD gives the average annual probability of default of the borrower, and EAD x LGD - the level of losses [6]. Thus, the expected losses can be estimated as follows:

$$EL = PD * EAD * LGD. \qquad (1)$$

Where EL— Expected Loss; PD - Probability of Default; EAD - Exposure at Default; LGD - Loss Given Default. symbols in your equation have been defined before or immediately following the equation.

The Basel Committee determined the following [12]:

● Average annual default probability (default rate) and borrower rating. PD is the probability that the loan cannot be repaid. The probability of default will be calculated for each employee. There are many different models that allow you to calculate PD based on available information. Three main categories can be distinguished: structural models, simplified models, and credit scoring models. The first two methods are based on market data (stock prices, bond yields). Therefore, the credit scoring model has the greatest practical significance. All value ranges are evaluated by the rating team. In addition, with the help of a special calibration of the rated value, the possibility of a default can be determined. PD, a similar rating group, has been elected president for one year [7];

●According to requirements exposure (EAD). EAD is an estimate of the amount of risk (that is, part of the loan). In this case, the following factors must be considered: First, the debt under the system (especially complex products with a limit system) may fluctuate over time, so it is necessary to evaluate its value in the event of default. [8]. Secondly, the existence of highly liquid collateral allows you to reduce EAD, so its implementation makes it possible to quickly repay part of the lost loan. However, the left part of the loan is unlikely to be repaid in full;

●The average expected loss of capital share (loss at default, LGD) at the time of default is usually calculated as a percentage of EAD. LGD is exactly the estimated loss of EAD. It is necessary to provide additional guarantees for the loan, assess the importance of the collateral to the customer and the current financial situation of the borrower, namely its rating. When calculating LGD and EAD, it is very important to correctly determine the issues of cost, liquidity and return probability.

Collect data on indicator evaluation and processing

The basis of the research is data on bank credit records. As the initial information, data describing credit history at the beginning of 2020 was selected based on monthly records of credit transaction cases. The following categories of credit history are introduced:

When analyzing a customer’s credit history, three main types of tasks are solved: comparing credit history indicators, analyzing the dynamics of credit operations, and analyzing the level of credit operations.

Table 1 - EL, PD, LGD, EAD results

Funded_amnt	PD	LGD	EAD	EL
5000	0.1647909316466200	0.9137287259284660	2949.6084487904	444.13495613990800
2500	0.2828104744734790	0.9154816249384510	1944.4333780507400	503.42895400336700
2400	0.2307362354090770	0.9194838219680120	1579.9343016863600	335.19607381040500
10000	0.20874299947855700	0.9049241978540790	6606.559611654160	1247.9565912633200
3000	0.1292812709740640	0.9114527959410170	2124.631666989780	250.3533717013720
5000	0.13580650461121300	0.9188613978157500	3174.1650557315200	396.095660557456
7000	0.29133725401367500	0.9165305685864820	5052.384586825450	1349.0852014910900
3000	0.2975299298493430	0.91369595859268	2056.324222637450	559.0156355556010
5600	0.41801183398990700	0.9002332778846770	4810.650248440880	1810.2869602668200
5375	0.1882600619740530	0.9198397777105950	3901.7992509115700	675.6710395256220
6500	0.180805842929102	0.9063090874655570	4788.668721435330	784.6998858150640

In order to collect data, it is necessary to select at least 50 customers belonging to the target industry group under consideration and collect the indicators that have been determined for them in the past few years. At the same time, data can be obtained from various sources (including those publicly available). In particular, data related to reports from the disclosure server, data from the Federal Office of State Statistics 2 etc. It is convenient to use related indicators (financial relations). In addition, financial indicators must be included in the same time range (quarter, month, year), and the duration of the production cycle must be considered in calculations.

Acknowledgment

Therefore, in modern banks, the role of risk management is significantly improving. When the decision to grant a loan is: accept/reject, the risk assessment allows you to give up the simple method. The credit risk management system becomes the basis for a reasonable dialogue with customers based on objectives including investment portfolio indicators. A properly structured business process will play an important role. The activities of the risk management department cannot be separated from other departments of the bank, because a correct and objective assessment requires all available information, including other departments and external resources. The role of risk management is not limited to the stage of considering the application - it is establishing the best portfolio structure, continuous monitoring, and the importance and understanding of existing risks.

It is impossible to completely eliminate the risk. There is no income without risk: the greater the risk, the higher the possible income. But you need to know the risks to prepare for possible consequences.

REFERENCES

1. V. Acharya. "Systematic Risk Theory and Prudent Banking Supervision Design", "Financial Stability Magazine", Volume 1. 5(3), 2007, 224-255 p.
2. E. Asarnow, D. Edwards. Measuring the loss of bank loans in arrears: a 24-year study. Journal of Business Loan, 2013, 11-23 p.
3. T. Bellotti, J. Crook. Model and predict the loss of credit card default settings. Working paper, Quantitative Financial Risk Management Center, 2007, 11-14 p.
4. S. Caselli, S. Gatti, F. Querci. The sensitivity of loss for a given default rate to systemic risk: new empirical evidence of bank loans. Journal of Financial Services Research, 34(1): 2008, 1-34 p.

5. M. Araten, M. Jacobs, P. Varshney. Measuring the default loss rate of commercial loans: an 18-year internal study. RMA Magazine, 2004, 96-103 p.
6. E. I. Altman. Default Recovery Rate and Default Loss Rate in Credit Risk Modeling and Practice: The latest review of literature and empirical evidence. Working paper, New York University Stern School of Business, 2006, - 123 p.
I. Schatz. The method of credit risk quantification: Probability of Default (PD). financial report. "BBVA Magazine", 2011, - 133 p.
7. S.G. Ryan, D.Vyas. Bank credit risk model disclosures have an impact on their loan loss preparation timeliness and loan origination. Working paper of New York University, 2012, 24-27 p.

Булантаев А.М.[1], Мұсахан Х.Б.[1], Молдагулова А.Н.[1], Сембина Г.К.[1]

Несие беру үшін банктен күтілетін шығындардың болжауы

Аңдатпа. Бұл мақалада статистикалық талдау мен банктер шығарған несиелер бойынша күтілетін шығындар болжамын енгізу үшін SAS платформасының үлгі деректері келтірілген. Осы зерттеудің бастапқы деректері банк клиенттерінің несиелік тарихы туралы ақпарат беретін Kaggle көзінен алынған. Технология логистикалық регрессияға, графикалық деректерді талдауға және SAS платформасында модель құру негіздеріне негізделген. Үлгіні несиелік тәуекелді болжау және банк жүйесіндегі несиелік тәуекелді сипаттау үшін пайдалануға болады.

Түйінді сөздер: компонент, деректерді талдау, несиелік тәуекел, берілген дефолттағы шығын, күтілетін шығын, дефолт ықтималдығы, дефолтқа ұшырау, логистикалық регрессия, модель, жұмыс істемейтін несиелер, арнайы әуе сервисінің платформасы, болжау.

Булантаев А.М.[1], Мұсахан Х.Б.[1], Молдагулова А.Н.[1], Сембина Г.К.[1]

Прогноз ожидаемых убытков банка при предоставлении кредита

Аннотация. В этой статье используются образцы данных платформы SAS в качестве примера для представления статистического анализа и прогнозирования ожидаемых убытков по кредитам, выданным банками. Исходные данные для этого исследования взяты из источника Kaggle, который предоставляет информацию о кредитной истории клиентов банка. Технология основана на логистической регрессии, графическом анализе данных и на основе построения модели на платформе SAS. Модель может использоваться для прогнозирования кредитного риска и описания кредитного риска в банковской системе.

Ключевые слова: компонент, анализ данных, кредитный риск, убыток при дефолте, ожидаемый убыток, вероятность дефолта, риск дефолта, логистическая регрессия, модель, неработающие займы, платформа Special Air Service, прогнозирование.

About the authors:

Bulantayev Aryslan Muratovich, second year master's degree student, Department of Information Systems, International Information Technology University.

Musakhan Khadisha Bakytzhankyzy, second year master's degree student, Department of Information Systems, International Information Technology University.

Moldagulova Aiman Nikolaevna, PhD in Physics and Mathematics, Associate Professor, Department of Information Systems, International Information Technology University.

Sembina Gulbakyt Kakeevna, PhD in Engineering Science, Associate Professor, Department of Information Systems, International Information Technology University.

Сведения об авторах:

Булантаев Арыслан Муратович, магистрант 2 курса кафедры «Информационных систем» Международного университета информационных технологий.

Мусахан Хадиша Бахытжанкызы, магистрант 2 курса кафедры «Информационных систем» Международного университета информационных технологий.

Молдагулова Айман Николаевна, к.ф.-м.н., доцент кафедры «Информационных систем» Международного университета информационных технологий.

Сембина Гулбакыт Какеевна, к.т.н., доцент кафедры «Информационных систем» Международного университета информационных технологий.

Авторлар туралы мәлімет:

Булантаев Арыслан Муратович, 2 курс магистранты, Халықаралық ақпараттық технологиялар университеті.

Мұсахан Хадиша Бақытжанқызы, 2 курс магистранты, Халықаралық ақпараттық технологиялар университеті.

Молдагулова Айман Николаевна, ф.-м.ғ.к., «Ақпараттық жүйелер» кафедрасының доценті, Халықаралық ақпараттық технологиялар университеті.

Сембина Гулбакыт Какеевна, т.ғ.к., «Ақпараттық жүйелер» кафедрасының доценті, Халықаралық ақпараттық технологиялар университеті.

УДК 336.77

Омарова Е.Г.

Казахстанско-Британский технический университет, Алматы, Казахстан

АЛГОРИТМ АВТОМАТИЗАЦИИ КЛАССИФИКАЦИИ ФИНАНСОВЫХ АКТИВОВ ПРИ РОЗНИЧНОМ КРЕДИТОВАНИИ

Аннотация. *В статье представлен кейс по составлению алгоритма автоматизации классификации финансовых активов при розничном кредитовании. Приведено заполнение спецификации основных бизнес требований, а также описан порядок разработки и внедрения в работу классификации. В табличном виде составлен Справочник пулов в виде портфеля однородных кредитов. Блок-схема иллюстрирует сущность предложений авторов по автоматизации процесса группировки розничных финансовых активов. С учетом технических особенностей и в целях реализации возможности расчета объема провизий по однородным кредитам, необходимо на каждом конкретном пуле завести поля.*

Ключевые слова*: финансовые активы, потребительское кредитование, розничные кредиты, однородные кредиты, определение провизий (резервов).*

Введение

Потребительское кредитование в Казахстане показывает взрывной рост. Как пишет журнал Forbes Kazakhstan и по сведениям Первого кредитного бюро, объем розничных кредитов на конец 2019 года составил 5,4 трлн тенге, что на 37,6% больше данных в 2018 году [1]. Банки наработали скоринговые системы кредитования. Автоматизации также подлежат финансовые активы при розничном кредитовании. В настоящей статье приведен пример составления алгоритма по автоматизации классификации финансовых активов с учетом розничных продуктов кредитования.

Отправной точкой взаимодействия структурных подразделений при инициации, оценке и реализации проекта в организации является согласование шаблонов спецификации бизнес-требований (Business Requirement Definition (BRD) к программному обеспечению. Понятие BRD — документа спецификации бизнес-требований введено Six Sigma [2]. Бесспорно, данный подход приемлем для всех заинтересованных и вовлеченных сторон в процессе проект-