МРНТИ 20.53.23
УДК 004.054

https://doi.org/10.51889/5184.2022.13.75.018

DEVELOPMENT OF SYSTEMS FOR EFFECTIVE ESTIMATION OF CREDIT SCORES

Aitim A.K.[1], Sembina G.K.[1]*

[1]International Information Technology University, Almaty, Kazakhstan
**e-mail: a.aitim@iitu.edu.kz*

Abstract

The article represented that in connection with the increasing importance of the time factor, automated methods of assessing creditworthiness are now becoming increasingly important. With mass consumer lending, practically in the presence of a potential borrower, it becomes necessary to decide on granting a loan. To operate profitably in the retail lending market, an effective risk assessment system is required, which would make it possible to cut off unreliable borrowers in advance and not refuse reliable borrowers, and reasonably determine the size of a consumer loan or a credit card limit. This system should create a margin of safety for the bank. One of such automated methods for assessing the creditworthiness of a counterparty is the scoring method.

Keywords: lending, credit scoring, score modelling, logistic regression, risk assessment system, lending market.

Аңдатпа

Ә.Қ. Әйтім [1], Г.К. Сембина [1]
[1]Халықаралық Ақпараттық Технологиялар Университеті, Алматы қ., Қазақстан

КРЕДИТТІК КӨРСЕТКІШТІ ТИІМДІ БАҒАЛАУ ЖҮЙЕЛЕРІН ӘЗІРЛЕУ

Мақалада уақыт факторының маңыздылығының артуына байланысты несие қабілеттілігін бағалаудың автоматтандырылған әдістері қазіргі уақытта барған сайын маңызды болып отырғаны көрсетілген. Несиені бағалау банктер несиені басқару шешімдерін қабылдаған кездегі маңызды үдерістердің бірі болып саналады. Жаппай тұтынушылық несиелендіру кезінде, іс жүзінде әлеуетті қарыз алушының қатысуымен несие беру туралы шешім қабылдау қажет болады. Несиелендіру нарығындағы табысты жұмыс істеу үшін сенімсіз қарыз алушыларды алдын ала кесіп тастауға және сенімді қарыз алушылардан бас тартпауға, тұтыну несиесінің сомасын немесе несие картасының лимитін негізді анықтауға мүмкіндік беретін тәуекелді бағалаудың тиімді жүйесі қажет. Бұл жүйе банк үшін қауіпсіздік маржасын құруы керек. Контрагенттің несиелік қабілетін бағалаудың автоматтандырылған әдістерінің бірі баллдық әдіс болып табылады.

Түйін сөздер: несиелеу, несиелік скоринг, скорингтік модельдеу, логистикалық регрессия, тәуекелді бағалау жүйесі, несие нарығы.

Аннотация

А.К. Айтим [1], Г.К. Сембина [1]*
[1]Международный Университет Информационных Технологий, г.Алматы, Казахстан

РАЗРАБОТКА СИСТЕМ ЭФФЕКТИВНОЙ ОЦЕНКИ КРЕДИТНЫХ РЕЙТИНГОВ

В статье представлено, что в связи с возрастанием значения фактора времени в настоящее время все большее значение приобретают автоматизированные методы оценки кредитоспособности. Оценка кредитоспособности считается одним из важнейших процессов, когда банки принимают решения по кредитному менеджменту. При массовом потребительском кредитовании практически в присутствии потенциального заемщика возникает необходимость принятия решения о выдаче кредита. Для прибыльной работы на рынке розничного кредитования необходима эффективная система оценки рисков, которая позволяла бы заранее отсечь неблагонадежных заемщиков и не отказать надежным заемщикам, обоснованно определить размер потребительского кредита или лимита по кредитной карте. Эта система должна создать запас прочности для банка. Одним из таких автоматизированных методов оценки кредитоспособности контрагента является метод скоринга.

Ключевые слова: кредитование, кредитный скоринг, скоринговое моделирование, логистическая регрессия, система оценки рисков, кредитный рынок.

1 Introduction

Credit evaluation is considered one of the most important processes when banks make decisions on credit management. This process involves the collection, testing and systematization of various credit components and variables for evaluating credit opinions. The quality of bank loans is considered the main factor determining the competitiveness, viability and profitability of banks. Credit scoring is considered one of the more powerful sets for systematizing bank buyers as part of the credit assessment process to reduce the current and expected risk of a bad credit situation of the buyer. To build a credit scoring model, you need to have high-quality data about borrowers, such as how the accuracy of the model depends on the data you choose to study it. Proper data is essential for predicting various types of lending, for example, for consumer loans and business loans, the models will differ [1].

For the subsequent construction of the model and assessment of its quality, it is important to consider the ratio of the amount of data corresponding to the default of customers and successful solvency. For the model to work correctly, it is necessary to have approximately the same number of both default and successful orders in the data. It is extremely important for the study to precisely determine the loan defaults, so there should be many such examples in the data.

2 Collection of Necessary Data and Initial Processing

The advances in technology have allowed money lenders to reduce credit risk by using all sorts of customer data. By applying statistical and machine learning techniques, cheap data is analyzed and reduced to a single value, known as a credit score, representing credit risk. This sense has the ability to assist in making a conclusion. The higher the credit rating, the more the lender is able to be in no doubt about the creditworthiness of the buyer. Credit scoring is an artificial intelligence configuration based on predictive modeling that considers the possibility that the customer will actually default on a loan promise, be delinquent or insolvent. The prediction model is “trained” by applying historical customer data along with peer group and other data to predict the possibility that a given customer will actually exhibit a particular behavior in the future [2].

The biggest advantage of credit scoring is the ability to quickly and effectively make decisions, for example, to accept or reject a buyer, or to increase or decrease the loan price, interest rate or term. As a result, the speed and accuracy of these judgments have made the credit rating the cornerstone of risk management across all sectors, spanning banking, telecommunications, insurance and retail.

Credit scoring can be used throughout the entire customer interaction cycle, including the duration of the customer interaction throughout the relationship between the customer and the organization. But they are primarily intended for credit risk departments, marketing departments still have every chance to benefit from credit scoring methods in their own advertising campaigns [3].

When evaluating orders, the risk of non-compliance with promises by fresh bidders is assessed when deciding whether to accept or reject orders. Behavioral evaluation considers the risk of default associated with an existing customer when making judgments regarding account management such as credit limit, overlimit management, fresh produce. Collection estimation is used in loan collection strategies to estimate the likelihood that buyers who dispose of the pledged asset will repay the obligation.

The development of a scientific and competent method and the prevention of individual proposals in this assessment using a credit rating system can be an effective step towards the optimal distribution of collected funds and the reduction of deferred receivables and, consequently, to improve the efficiency of the banking system. Satisfactorily, credit scoring models increase the efficiency of credit solutions in the production of services and meeting the needs of customers and will also be able to reduce the causes of material needs and default of borrowers. Therefore, it is necessary to develop a calming tool for measuring the credit risk of its customers, and for this tool it is necessary to do qualitative and quantitative credit risk using the credit scoring method.

The exchange of information on the characteristics of loans for loans and the proportion of their debt can have a significant impact on the effectiveness of credit markets.

- Firstly, the exchange of information and knowledge improves the bank's understanding of the characteristics of applicants for benefits and gives more accurate forecasts of the relative probability of placement.
- Secondly, according to the fundamentals of credit alienation of banks, the interest rate on their credit resources may fluctuate depending on the level of risk of applicants to a certain extent with the strengthening of monetary and tax policy.
- Thirdly, this system can be used to create loan recipients, and, after that, the indicator can reduce the motivation of bank customers to receive additional services and exceed their capabilities by placing various banks and modeling their own condition (without expressing and declaring the real balance of their own received funds) [4].

The Process of Building a Valuation Map Model

Over the years, several different modeling techniques have been observed to implement credit scoring. They range from parametric or non-parametric, statistical or machine learning to supervised or unsupervised algorithms. the latter methods involve rather difficult spreads involving hundreds or thousands of different models, different test structures, and ensemble methods with several learning methods to achieve greater accuracy.

Ignoring this diversity, the method of modeling is different - the credit rating model. Commonly referred to as a normal scorecard, it is based on logistic regression as the underlying model. In comparison with other modeling methods, this method meets almost all the claims, in fact, which prepares it with the desired alignment between practitioners and is used by almost 90% of the creators of indicator charts. A feature system model is simple to build, understand, and implement, and can be nimbly implemented. Being a hybrid of statistics and machine learning, its predictive accuracy is comparable to other more complex methods, and its estimates have every chance of being applied precisely as probability estimates and, therefore, for direct data entry for risk-based pricing. This is quite fundamental for lenders that comply with the Basel II regulatory framework. While instinctive and straightforward to interpret and justify, scorecards are mandated by regulators as the exclusive way to model credit risk in some countries.

The business intelligence test uses these methods as a measurement data test, a method of combining univariate and multivariate statistics, and all kinds of data visualization methods [5]. Correlation, cross-tabulation, scatter, timeline test, and supervised and unsupervised segmentation test are considered common methods. Segmentation is special because it determines when a certain number of scorecards are needed.

The choice of variables based on the results of business intelligence analysis is introduced with the division of the data mining view as a minimum number into 2 different sections: the section for study and testing. Variable selection is a set of model candidate variables whose significance is tested during model learning. Candidate model variables are still popular as autonomous variables, predictors, attributes, model moments, covariates, regressors, functions, or properties [6].

The main task is to find the right set of variables so that the scorecard model can not only rank buyers based on the likelihood of their having bad debts, but also consider the possibility of their having bad debts. Typically, this means choosing statistically important variables in the predictive model and having an equilibrium set of predictors (usually 8-15 is a good balance) to approximate a 360-degree view of buyers. In addition to buyer-specific risk features, we still need to consider the likelihood of connecting regular risk moments to account for financial drift and volatility [7].

When choosing variables, there are a few restrictions:

- For starters, the model typically has some high predictor variables that are prohibited by legal, ethical, or regulatory rules.
- Secondly, some variables have every chance of being unattainable or of low quality at the modeling or manufacturing steps. Apart from this, there is every chance of being significant variables that were not recognized as such, for example, due to the periodic oversight of the population selection or because of such that their model effect would be inconsistent due to multicollinearity.
- And, in the end, the last text every time will be the case, and he can insist on connecting only variables that are significant for the business, or to insistently ask for uniformly growing or decreasing effects.

All these limitations are considered likely sources of periodic misses, which in fact makes it difficult for data scientists to minimize periodic selection misses. Common preventive measures in variable selection include:

- Collaboration with experts in the given field to identify significant variables;
- Awareness of every dilemma related to data source, reliability or measurement error;
- data cleaning;
- Introduction of control variables to account for unresolved variables or certain activities such as financial drift. It is important to understand that the choice of variables is an iterative process that happens throughout the entire process of building a model.
- It occurs before model fitting by reducing the number of variables in the data mining view to a manageable set of candidate variables;
- lasts in the process of learning the model, where subsequent reduction is performed as a result of statistical insignificance, multicollinearity, small contributions, or penalties to avoid overfitting;
- Lasts during the evaluation and testing of the model;

- Ends during business assertion, where the readability and interpretability of the model play a significant role [8].

The selection of variables is completed after reaching the "golden spot" - this means that further improvement in terms of model accuracy cannot be achieved. The iterative nature of the variable selection process is shown in Fig.1.

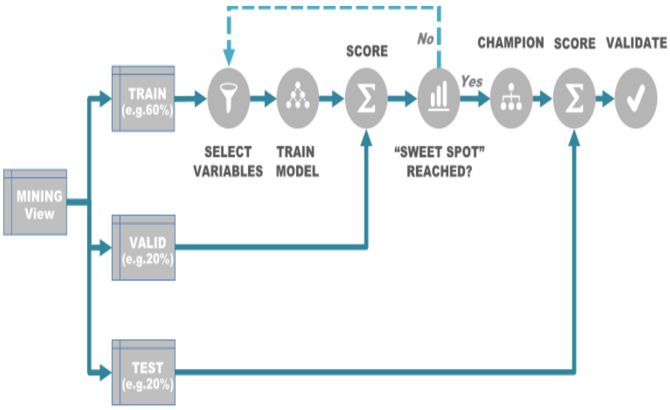

Figure 1. Iterative nature of the variable selection process

In addition to some general recommendations for solving this problem, the data specialist should offer the best approach to converting the signature of the client's data into a powerful information artifact - a representation of data mining. This is probably the most creative and most challenging aspect of the data scientist role, as it requires a solid understanding of the business in addition to statistical and analytical skills. Very often, the key to creating a good model is not the power of a particular modeling method, but the breadth and depth of derived variables that represent a higher level of knowledge about the phenomena being studied.

3 Preparation and Development of a System Indicators

The data preparation process begins with data collection, commonly referred to as an ETL (extract-transform-load) move. Data integration brings all kinds of informants together using data joining and grouping. As a rule, this requires the manipulation of relational tables with the implementation of several rules of unity, such as entity unity, referential unity, and domain unity [9]. Applying one-to-one, one-to-many, or many-to-many cases, the data is aggregated to an important analysis value, resulting in the original signature of the buyer. The process of preparing data for filling in the scorecard is shown in Figure 2.

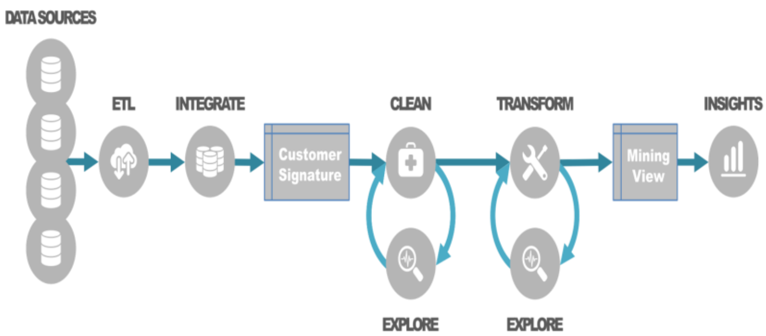

Figure 2. Data preparation process

Before deciding how to cultivate missing meanings, we need to understand the basis of missing data and understand the distribution of missing data so that we can systematize them as:

- Completely absent by accident (MCAR);
- Missing by accident (MAR);
- Missing is not accidental (MNAR).

Handling missing data is often associated with MCAR and MAR, during which time it is more difficult to work with MNAR.

The presence of outliers has the potential to fail the statistical assumptions on which we intend to build the model. Subsequently identifying is fundamentally to understand the background of the outliers before using any kind of healing. For example, outliers have every chance of being a valuable source of information when fraud is detected; as a result, it would be a bad idea to change them with the mean or median meaning.

Data mining and data cleansing are considered mutually cyclical steps Data mining includes both univariates, eg, and bivariate testing and ranges from univariate statistics and frequency spreads to correlations, crosstabs, and data analysis. A univariate exploratory data test is shown in Figure 3.

Univariate View

Variable	NMiss	Min	Max	Mean	Std Dev	Variance	Kurtosis	Skewness	2.5%	97.5%
age	0	17	90	38.58165	13.64043	186.0614	-0.16613	0.558743	18	68
education_num	0	1	16	10.08068	2.57272	6.61889	0.623444	-0.31168	4	15
capital_gain	0	0	99999	1077.649	7385.292	54542539	154.7994	11.95385	0	8614
capital_loss	0	0	4356	87.30383	402.9602	162376.9	20.3768	4.594629	0	1887
hours_per_week	0	1	99	40.43746	12.34743	152.459	2.916687	0.227643	12	66
ID	0	1	32561	16281	9399.695	88354274	-1.2	-457E-20	815	31747
churn	0	0	1	0.24081	0.427581	0.182826	-0.53005	1.21243	0	1
NUM_TRANS	197	1	17	5.030435	2.202696	4.851869	0.151042	0.487854	1	10
MAX_AMOUNT	197	7	5000	4031.597	914.8534	836956.7	2.158614	-1.48581	1556	4974
MIN_AMOUNT	197	0	4982	967.4932	919.1578	844851.1	2.279718	1.515301	25	3521
AVG_AMOUNT	197	7	4982	2500.48	739.0471	546190.7	0.342705	0.020696	1015.833	4020.6
SUM_AMOUNT	197	7	46553	12577.97	6396.807	40919141	0.306289	0.594045	2143	26534
RECENT	197	1	11	2.656934	1.973419	3.894384	2.30203	1.531862	1	8

Summary View | Data | Univariate View | Univariate Charts | Predictive Power

Figure 3. EDA (one-dimensional view)

Subsequently, exploratory data analysis (EDA) data is processed to increase properties [10]. Data cleansing requires good business conduct and data awareness so that the data can be correctly interpreted. It is an iterative process designed to eliminate violations and replace, reconfigure, or remove these violations as needed. The 2 main difficulties with dirty data are missing meanings and outliers; both have every chance of strongly influencing the accuracy of the model, because of which prudent intervention is needed.

Variable Transformations

Scorecard development outlines how to turn the data into a scorecard model if data preparation and the initial variable selection process (filtering) are completed and the filtered training dataset is available for the model building process. The development process is made up of 4 main parts: rearranging the variables, learning the model with the implementation of logistic regression, testing the model and scaling. Figure 4 illustrates the process of developing a feature system.

The usual scorecard model based on logistic regression is seen as an additive model. Accordingly, special rearrangements of variables will be required.

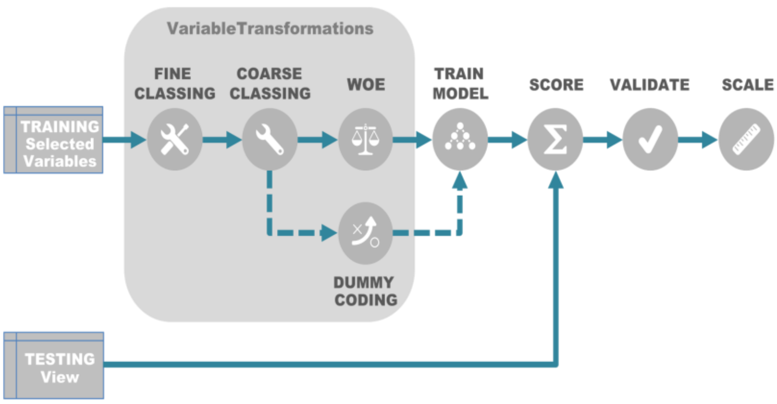

Figure 4. The process of developing a system of indicators

Training, Scaling and Evaluation of the Model

Model evaluation is considered the last step in the model construction process. It is produced from 3 separate milestones: evaluation, testing, and acceptance.

The main metrics assessed are statistical characteristics, covering model accuracy, complexity, miss rate, model correlation statistics, variable statistics, sense of significance, and odds ratios [11].

The choice of test metric depends on the similarity of the model classifier. The most common indicators for binary systematization problems are the lifting diagram, the lifting force diagram, the ROC curve, and the Kolmogorov-Smirnov diagram. The ROC curve is the most common inventory for visualizing model data. This is a universal tool that is used for:

- champion-challenger methodology for choosing a more efficient model;
- Testing the performance of the model on invisible data and comparing it with the training data;
- Choosing a rational threshold that maximizes the number of true positives and minimizes the number of false positives.

The ROC-curve is based on the method of raising the dependence of sensitivity on the probability of false positives (false positives ratio) at all possible thresholds. A desirable feature of the ROC curve is the assessment of performance characteristics at various thresholds. Different types of business problems will have different threshold meanings depending on the business strategy.

The area under the ROC curve (AUC) is a useful indicator to indicate the predictive ability of a classifier. As far as credit risk is concerned, an AUC of 0.75 or higher is considered an industry stereotype and a sine qua non for model acceptance. Figure 6 shows the performance characteristics of the model.

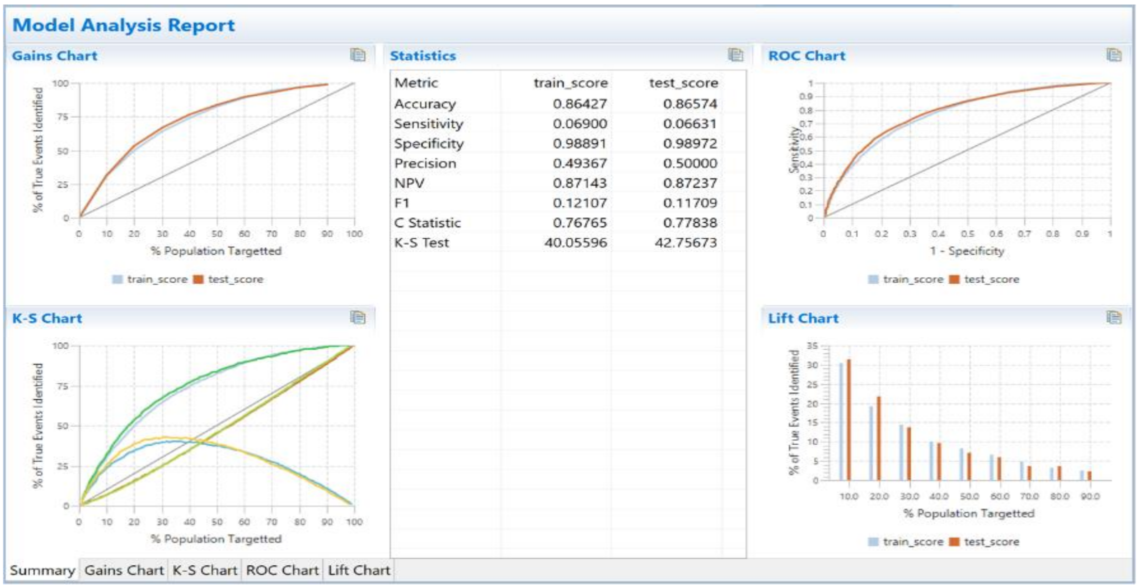

Figure 5. Performance indicators of the model

Receiving utility is seen as a critical boundary when a data professional is obliged to rebuild models for business and “protect” this model. The main aspect of the evaluation is the financial benefit of the model, because of which the benefit test occupies a central place in the presentation of the results. Data scientists are required to make every effort to suggest summaries in short form so that the cross section and output are simply skipped and understood. Failure to receive this can lead to withdrawal from the model and, therefore, to a breakthrough of the plan.

Conclusion

Banks face a wide range of risks in their day-to-day operations. The main activity of banks is to raise funds by issuing various loans to individuals and legal entities. The subsequent assessment of credit risk is one of the main tasks in the banking sector. The production methods used by banks to create scoring models are analyzed. Various statistical approaches to the analysis of the model quality are considered. Automation of such a routine procedure as evaluation allows banks to reduce the costs of joint operations, freeing up labor and financial resources to do other tasks. Therefore, the use of modern machine learning algorithms can help in solving real business problems.

References:

1 Gulzat T., Lyazat N., Siladi V., Gulbakyt S., Maksatbek S., (2017) Research on predictive model based on classification with parameters of optimization, Neural Network World, 30 (5), pp.295-301.

2 Bulantayev A.M., Musakhan K.B., Moldagulova A.N., Sembina G.K., (2016) "Forecasting expected bank losses at granting a loan," International Journal of Information and Communication Technologies, pp.154-159.

3 Satybaldiyeva R., Uskenbayeva R., Moldagulova A., Kalpeyeva Z., and Aitim A., (2019) "Features of Administrative and Management Processes Modeling", World Congress on Global Optimization, pp. 842-849, doi:10.1007/978-3-030-21803-4_84.

4 Lee T.S., and Chen I.F., (2015) "A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines", Expert Systems with Applications, 28(4), pp. 743-752, doi:10.1016/j.eswa.2004.12.031.

5 Edelman D.B., and Crook J.N., (2016) "Credit scoring and its applications. Society for Industrial Mathematics", pp. 184-192.

6 Thomas L.C., (2017) "A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers", International Journal of Forecasting, 16(2), pp. 149-172, doi:10.1016/S0169-2070(00)00034-0.

7 Abdou H.A., and Pointon J., (2015) "Credit scoring, statistical techniques and evaluation criteria: a review of the literature", Intelligent Systems in Accounting, Finance and Management, pp. 45-66.

8 Leung K., (2015) "A comparison of variable selection techniques for credit scoring", pp. 53-59.

9 Tsai C.F., and Chen M.-L. , (2018) "Credit rating by hybrid machine learning techniques", Applied Soft Computing, 10(2), pp. 374-380, doi:10.1002/isaf.325.

10 Luo S.T., Cheng B.-W., and Hsieh C.-H., (2017) "Prediction model building with clustering-launched classification and support vector machines in credit scoring", Expert Systems with Applications, 36(4), pp. 562-566.

11 Nanni, L. and A. Lumini, (2015) "An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring". Expert Systems with Applications, 36(2), pp. 302-303.