European Proceedings of
Social and Behavioural Sciences
EpSBS

www.europeanproceedings.com
e-ISSN: 2357-1330

DOI: 10.15405/epsbs.2020.03.109

II International Scientific Conference GCPMED 2019
"Global Challenges and Prospects of the Modern Economic Development"

MACHINE LEARNING MODELS FOR PREDICTING BANKRUPTCY OF ENTERPRISES

A. N. Opekunov (a), N. V. Nikitina (b)*, N. V. Kamardina (c)
*Corresponding author

(a) Penza State University, Russia, 440026, t. Penza,str. Krasnaya, 40, Opekunovan@yandex.ru
(b) Samara State University of Economics, Russia, Samara, Soviet Army str., 141, 443090, nikitina_nv@mail.ru
(c) Penza State University, Russia, 440026, t. Penza, str. Krasnaya, 40, alias-melafon@mail.ru

Abstract

In the article is given the existing models estimation and classification of forecasting of enterprises bankruptcy. Most of the models used in practice are poorly adapted to Russian conditions, and models of domestic development do not meet the requirements of adequacy and accuracy of estimates. The necessity of forming a new methodological approach to the development of prognostic models, based on the principles of machine learning, by which we understand the process of choosing the optimal forecasting model, based on minimizing the functional error in assessing the quality of the selected model. The implementation of this approach is considered on the example of Logit-model built on the materials of enterprises formed in the training sample. The training sample was compiled on the data of financial statements of enterprises that entered bankruptcy proceedings, as well as retained solvency in one time period. In order to predict bankruptcy, it is proposed to use models with binary classification training, which allow not only to estimate the probability of occurrence of an event, but also to classify the objects of study. It is proved that it is impossible to build models of binary classification of linear type. In addition, Logit models are more accurate and allow you to include not only quantitative but also qualitative factors in the model. The model of forecasting bankruptcy of the enterprises developed by this technique showed high accuracy on a training sample-85% of accuracy of an assessment. The expediency of further research on this topic is proved.

Keywords: Bankruptcy, forecast, model, machine learning.

https://doi.org/10.15405/epsbs.2020.03.109
Corresponding Author: N. V. Nikitina
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

1. Introduction

Instability, low growth rates of the Russian economy are accompanied by sharp and periodic bursts of bankruptcies of enterprises in various sectors. At the same time, there is a shortage of scientific and practical research explaining the causes and consequences of this problem. Most anti-crisis measures are macroeconomic in nature, avoiding the problems of specific economic units. At the same time, the achievement of targets for sustainable development and growth of the country's economy, in our opinion, will be impossible without a scientifically sound solution to this problem.

2. Problem Statement

The problems of forecasting enterprises bankruptcy are considered today in many scientific papers of both Russian and foreign authors. However, it should be noted that most foreign methods are poorly adapted to Russian conditions. At the same time, forecasting models based on domestic methods do not always correspond to the declared accuracy and adequacy (Kopelev, 2014). There is a high need to develop methods for predicting the bankruptcy of enterprises, which would be characterized by high accuracy, flexibility of models and, at the same time, moderate complexity of their construction.

In modern conditions there are new tools of economic and mathematical modeling-data analysis and machine learning, allowing to solve such problems. The implementation of the principles of machine learning allows not only to build predictive models, but also to effectively adapt them to the rapidly changing conditions of the market environment (Malkina & Ovcharov, 2019).

3. Research Questions

Despite the huge variety of proposed models for predicting bankruptcy of enterprises, they can be narrowly classified. Thus, most models are built on the principles of linear regression, where financial coefficients act as regressors. Such approaches include models of Beaver, Fox, Tishou, etc. (Lacombe & McIntyre, 2017).

In the models of Altman, Fulmer, Springate, the elements of discriminant analysis are applied, which allows to effectively distribute the objects of study by classes-to rank according to the degree of probability of bankruptcy (Bauer & Agarwal, 2014).

Most domestic models of bankruptcy forecasting are formed on the basis of these two approaches. These include: Zaitseva model, model of Saifullin and Kadykov, model of Irkutsk GEA, Kolyshkin model and others (Kazakova, 2019). Recently, models based on logit regressions have become more common. The use of logit models allows for higher prediction accuracy by including both quantitative and qualitative variables in the model (Brîndescu-Olariu, 2017). The most famous logit models are Olson, Evstropov, Haidarshina (Shchepot'ev, 2019).

Modern trends in the development of scientific tools in the economy and form new approaches to the development of models for predicting bankruptcy, based on the use of modern economic and mathematical methods (machine learning models, neural network models, artificial intelligence models (Fedorova & Timofeev, 2015).

https://doi.org/10.15405/epsbs.2020.03.109
Corresponding Author: N. V. Nikitina
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

4. Purpose of the Study

Unstable dynamics of economic processes, avalanche growth of information resources on changes in various factors of enterprises determines the need to develop and implement new approaches to the formation of models for forecasting bankruptcies.

5. Research Methods

According to the results of the analysis of methods for predicting bankruptcy of enterprises, it is determined that the most commonly used models for this purpose are based on linear and logit regressions. Moreover, logit models have a higher predictive ability. Based on this, we will form the elements of the methodology for constructing a logit model for predicting bankruptcy of enterprises, using the principles of machine learning.

In problems with training it is necessary to form a forecast of discrete feature Y, knowing a set of values of exogenous factors x1, x2,..., xn. For the purposes of bankruptcy forecasting, in our opinion, it is advisable to use binary classification machine learning models (Mohri, Rostamizadeh, & Talwalkar, 2012). In this approach, the dependent variable or binary attribute Y takes two values:

- 1 if the event occurs;
- 0 if the event does not occur.

The use of linear regression models in binary classification problems is impossible, because the dependent variable Y here takes any values from -∞ to +∞. The process of finding and constructing the optimal function predicting the probability of a is called learning. At the heart of this process is the minimization of the quality functional (error), which measures the quality of the model (Keller, Kim, & Steiner, 2013). Using the logit regression model, we analyze the dependence of bankruptcy on the value of the current liquidity ratio. We calculated the values of feature x-the current liquidity ratio and the value of feature Y, equal to zero for those firms that are successfully functioning, and one for those that have become bankrupt, for twenty firms (table 01).

Table 01. Estimation of logit-model parameters and verification of the hypothesis of its significance

x	**y**	**a0**	**7,145969**	**IN L**	**-5,739361961**
0,880042852	1	a1	-6,31325	In L (при =const)	-12,47664925
1,099836572	1				
0,133160244	1	p^	0,5	Likelihood ratio	13,47457458
0,556478261	1			Number of degrees of freedom	1
0,088398018	1			P-value	0,000241818
0,259866144	1				
1,376578862	1			x	1,5
1,004014513	1			P^{y=1|x}	0,089162507
0,778048468	1				
0,424361942	1				
1,672766965	0				
1,421899886	0				
1,876226089	0				
1,018889339	0				
1,389917326	0				
1,080869776	0				
3,439537102	0				
1,267136327	0				
9,229677076	0				
8,359749012	0				

https://doi.org/10.15405/epsbs.2020.03.109
Corresponding Author: N. V. Nikitina
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

Parameters are evaluated in the selected cells: a_0 = 7,145969 и a_1 = -6,31325.

Thus, the model has the form: $ln\frac{p}{1-p} = a_0 + a_1 x$ =7,145969-6,31325x

or

$$p = \frac{1}{1 + e^{-7,145969+6,31325x}}$$

Next, we define the likelihood ratio. Since there is only one regressor in the model, we introduce the number of degrees of freedom equal to one and calculate the observed significance level of the hypothesis H0: a1 = 0. Since it turned out to be equal to 0.000241818, there is reason to reject the hypothesis H0 and assume that the probability of bankruptcy depends on the current liquidity ratio at any significance level greater than 0.000241818.

A negative value of the coefficient a1 indicates that the smaller the CTL, the higher the probability that the firm will go bankrupt. More precisely, the decrease in the current liquidity ratio by 0.01 points is accompanied by an average increase in the chance of bankruptcy by e^(-6.31325)=0.0018 times.

6. Findings

On the Figure 01, blue dots indicate observations, and square boxes indicate predictions obtained by clipping (the firm is considered successful at predicting probability p(x) ≥ 1.5 and bankrupt at p (x) < 1.5.

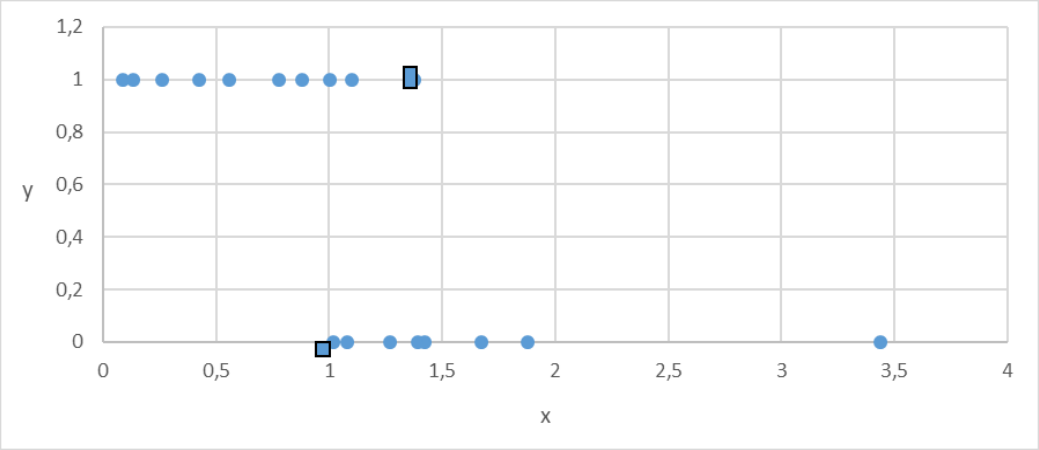

Figure 01. Estimation of accuracy of the predictive model on the learning sample [author's calculation]

You can see that three blue dots are circled: two bankrupt companies are recognized as successful, and one successful-bankrupt. Overall, 17 of the 20 observations were correctly classified and 3 of the 20 were incorrectly classified. Thus, the proportion of correct forecasts was 85%.

https://doi.org/10.15405/epsbs.2020.03.109
Corresponding Author: N. V. Nikitina
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

7. Conclusion

In General, the proposed method using machine learning elements can be used to develop models for predicting bankruptcy of enterprises. The model developed by this technique already gives a high degree of accuracy of the forecast on the training sample. However, for the purpose of its real testing, it is necessary to expand the data set on the training sample both by the number of enterprises studied and by the justified need to include additional factors in the forecast model.

References

Bauer, J., & Agarwal, V. (2014). Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking & Finance, 40*(C), 432-442. https://doi.org/10.1016/j.jbankfin.2013.12.013

Brîndescu-Olariu, D. (2017). Bankruptcy prediction logit model developed on Romanian paired sample. *Theoretical and Applied Economics, XXIV*(1(610)), 5-22.

Fedorova, E. A., & Timofeev, Ya. V. (2015). Developing the bankruptcy prediction models for Russian businesses of the construction and agriculture industries. *Finances and Credit, 21*(32), 2-10. [in Rus.].

Kazakova, N. A. (2019). A risk-oriented model to assess the probability of bankruptcy of strategic alliance member companies. *Economic Analysis: Theory and Practice, 18*(7), 1295-1308. https://doi.org/10.24891/ea.18.7.1295 [in Rus.].

Keller, B. S. B., Kim, J. -S., & Steiner, P. M. (2013). Abstract: Data mining alternatives to logistic regression for prope. *Multivariate Behavioral Research, 48*(1), 164. https://doi.org/10.1080/00273171.2013.752263

Kopelev, I. B. (2014). Factors of predicting bankruptcy risk of a company. *Finances and Credit, 20*(45), 42-49. [in Rus.].

Lacombe, D. J., & McIntyre, S. G. (2017). Hierarchical spatial econometric models in regional science. In R. Jackson, & P. Schaeffer (Eds.), *Regional Research Frontiers Methodological Advances, Regional Systems Modeling and Open Sciences, 2*(17) (pp.151-167). Cham: Springer. https://doi.org/10.1007/978-3-319-50590-9_9

Malkina, M. Yu., & Ovcharov, A. O. (2019). Assessing the financial instability of economic systems: A variety of methods and models. *Economic Analysis: Theory and Practice, 18*(7), 1273-1294. https://doi.org/10.24891/ea.18.7.1273 [in Rus.].

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2012). *Foundations of machine learning.* Cambridge, Massachusetts: MIT Press.

Shchepot'ev, A. V. (2019). Specifics of determining the value of net assets and other financial indicators in the financial analysis of entities going bankrupt. *Economic Analysis: Theory and Practice*, *18*(5), 974-994. https://doi.org/ 10.24891/ea.18.5.974 [in Rus.].