ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-6-131-155
UDC 336.6(045)
JEL G33, G17

Bankruptcy Risk Factors of Russian Companies

A.A. Zhukov[a], E.D. Nikulin[b], D.A. Shchuchkin[c]
[a, b] St. Petersburg State University, St. Petersburg, Russia;
[c] Bonch-Bruevich St. Petersburg State University of Telecommunications, St. Petersburg, Russia

ABSTRACT

The bankruptcy of Russian companies in the existing environment has become rather common. Determination of bankruptcy risk factors allows predicting the prospects for business development. The authors set the task to determine the relative influence of individual financial and non-financial factors on the probability of a company's bankruptcy. To study risk factors, the authors analyzed 3184 large Russian companies (with revenues of more than 2 billion rubles per year and more than 250 employees) of various industries operating from 2009 to 2020. The total number of observations is 38,208. For analysis, 30 factors were selected and divided into five groups: profitability, liquidity, turnover, financial stability and general (non-financial) factors. For the study, one of the machine learning methods was used — the random forest method. The sample consists of companies from seven industries, including manufacturing, retail, construction, electric power, mining, agricultural production, and water supply, as well as other industries, which include companies in education, healthcare, agriculture, and hospitality. The analysis was carried out both in aggregate for the entire sample without being distributed by industry, and for samples distributed by manufacturing, retail, and service industries. In the sample as a whole, the tested model in 86% of cases correctly predicted the possibility of a company going bankrupt for the period under review. This result confirmed that machine learning methods (in particular, the random forest algorithm) are highly effective in solving the problem of bankruptcy prediction for a company. Based on the data obtained, the paper concludes that profitability factors have the most significant impact on the probability of bankruptcy for manufacturing and retail companies. For service companies, it is financial stability factors. Solving the problem of determining the bankruptcy risk factors of Russian companies will ensure a reduction in the number of bankrupt enterprises, which, in turn, will contribute to the recovery and development of the national economy.
Keywords: corporate finance; large companies; business; financial analysis; financial stability; bankruptcy prediction; bankruptcy risk factors; machine learning methods; random forest

For citation: **Zhukov A.A., Nikulin E.D., Shchuchkin D.A. Bankruptcy risk factors of Russian companies. *Finance: Theory and Practice*. 2022;26(6):131-155. DOI: 10.26794/2587-5671-2022-26-6-131-155**

INTRODUCTION

The problem of evaluating the performance of companies, as well as forecasting the prospects for business development, is acute for various economic agents. It is important for internal and external users of accounting information to understand what indicators should be focused on in the first place in the financial analysis of a business in order to predict bankruptcy risk.[1] For the Russian market, such an analysis is especially relevant. According to the Center for Macroeconomic Analysis and Short-Term Forecasting (CMASF), since 2016 there has been a steady downward trend in the number of operating organizations in the Russian economy.[2] According to the report of the Central Bank of Russia, the maximum number of bankruptcy reports since 2015 was reached in 2017, exceeding 11,000.[3]

One of the first mathematical models for assessing the bankruptcy risks of companies that have become widespread in practice is the model of E. Altman, proposed in 1968 [1]. The first version of the Altman model was formulated using the method of *multiple discriminant analysis* of 66 American companies, some of which went bankrupt during the observation period, and some continued to operate. The main advantage of the model, which determined its practical significance, is the integral indicator of the financial condition of an enterprise (*Z-score*), developed on its basis, which allows ranking organizations according to the degree of risk of bankruptcy. In later publications, a cross-country analysis of the E. Altman model was carried out and the applicability of the model for various markets was determined [2, 3]. Many models for predicting the bankruptcy of a company that appeared in the 1970s-1980s were methodologically based on the Altman model (see, for example, the model of R. Taffler [4]).

An important stage in the development of models for predicting the bankruptcy of a company was the works of M. Zmijewski [5] and J. Ohlson [6] published in the 1980s. In the M. Zmijewski's work, the probit model was used, and Ohlson's approach was based on the logit model. The method of logistic regression is still used in many domestic and foreign studies on this issue (see, for example, [7–10]).

Modern studies of company bankruptcy are based on advanced methods of statistical data analysis, primarily on *machine learning* methods [11–15]. The main reason for the spread of these methods in the analyses of the risk of bankruptcy of companies is that they allow overcoming the shortcomings of regression models, which are expressed, in particular, in a decrease in the predictive power of these models in the case of a non-linear relationship between variables [7].

Macroeconomic features of country markets, as well as differences in current legislation and accounting standards, inevitably require the adaptation of foreign bankruptcy prediction models to the specifics of a particular country and/or the development of original models based on data of companies in this country. A systematic analysis of the literature conducted in 2018 by A.V. Kazakov and A.V. Kolyshkin, based on more than 40 domestic bankruptcy prediction models [16], demonstrated an average "low" quality of research in this field. The shortcomings of the existing studies, in particular, include the small sample size on which the models' parameters were estimated and the lack of a test sample. A small number of publications about Russian companies

[1] In accordance with Art. 2 of the Bankruptcy Law of the Russian Federation, insolvency (bankruptcy) is "a debtor's inability recognized by an arbitration court to fully satisfy the claims of creditors for monetary obligations, for the payment of severance benefits and (or) for remuneration of persons working or working under an employment contract, and (or) fulfill the obligation to make mandatory payments", the Federal Tax Service of the Russian Federation. URL: https://www.nalog.gov.ru/rn77/taxation/bankruptcy/ (accessed on 17.02.2021).

[2] Fundamental Research Program of the National Research University Higher School of Economics. The bankruptcy of legal entities in Russia: Main trends. 19.01.2021. URL: http://www.forecast.ru/default.aspx (accessed on 25.05.2021).

[3] Monetary Policy Report of the Central Bank of the Russian Federation, July 2021. URL: https://cbr.ru/analytics/dkp/ddcp/longread_3_35/page/ (accessed on 05.04.2021).

that would use modern methods of machine learning should also be noted [7]. Thus, the relevance of research on the issues of the bankruptcy predictions of Russian companies remains.

This study differs from the works existing on the Russian market, in which the bankruptcy risks are analyzed using machine learning methods, in the following parameters. *First*, the focus of most studies of this kind (see, for example, [7, 17–19]) is the problem of comparing the *predictive power* of different models, usually regression models and models based on machine learning. At the same time, this paper aims to address another problem — it determines the *relative influence* of individual financial and non-financial *factors* on the bankruptcy probability of a company. The analysis carried out by the random forest method allows us to rank the considered indicators according to the degree of priority in assessing the bankruptcy risk of the company. *Second,* existing studies either do not imply a comparative intersectoral analysis (i.e., the analysis is carried out on the entire sample) [18, 19], or they study only one industry [7, 17]. In this paper, the analysis is carried out both for the entire sample and for industry subsamples —manufacturing, retail, and service industries. *Third,* an important result of this paper is the determination of the ranges of values of the corresponding indicators characterizing the different degrees of risk of bankruptcy of an enterprise, namely "high", "medium" and "low". This result makes it possible to classify companies in a form convenient for practical use, depending on the degree of probability of their bankruptcy.

Thus, the purpose of this study is to identify the main risk factors for the bankruptcy of Russian companies.

The authors analyzed large Russian companies (with revenues of more than 2 billion rubles a year and more than 250 employees) of various industries. The observation period is 12 years (from 2009 to 2020). The sample included data on 3184 companies, i.e. the total number of observations is 38,208. The main method of analysis is one of the machine learning methods — *the random forest*. More than 30 indicators were analyzed as potential factors of bankruptcy, including non-financial ones (for example, the number of employees or the form of ownership). Financial indicators (factors) were divided into four groups: liquidity, profitability, financial stability, and turnover.

MACHINE LEARNING METHODS IN BANKRUPTCY RESEARCH OF RUSSIAN COMPANIES

In recent years, in research on the analysis of the bankruptcy risk of a company, advanced methods of statistical data analysis, primarily machine learning methods, are increasingly used. Their active use in studies of Russian companies dates back to the 2010s.

The main result of most of the studies carried out is the conclusion that machine learning methods make it possible to obtain more accurate predictions of the company's bankruptcy probability compared to traditional methods of data regression analysis. For example, the study by B.B. Demeshev and A.S. Tikhonova, conducted in 2014, compared the predictive power of various statistical methods, including both traditional and more modern approaches: logit and probit models; models based on discriminant analysis; classification trees and random forest. The sample was limited to data collected from medium and small businesses in 2011–2012. According to the research results, it turned out that the most accurate tool for predicting the bankruptcy of a company is such a machine learning method as a random forest [20].

The study [17] also concluded that machine learning methods are characterized by higher accuracy in predicting the probability of bankruptcy. Its authors, having analyzed 5120 Russian companies in the production and distribution of electricity, gas, and water for 2009, argue that the most accurate tool for predicting bankruptcy is a neural network, which showed a higher percentage of correct bankruptcy predictions compared to discriminant analysis and a logit model.

A similar conclusion was made in the study by A.M. Karminsky and R.N. Burekhin [7]. The authors compared the predictive power of a large number of models, including various neural network modifications, classification trees, and random forests. The analysis of companies in the Russian construction industry for 2011–2017 showed that in the sample under review, the best results are obtained by an artificial neural network with one hidden layer and four neurons.

In general, studies show that machine learning methods in general and neural networks in particular can significantly increase the efficiency of company bankruptcy risk analysis, but the effectiveness of these methods depends on the quality and availability of the initial data [21–23].

A number of scientific papers are devoted to the development of new algorithms/models for predicting the financial insolvency of Russian companies based on machine learning methods. For example, the study [23] analyzed the financial stability of manufacturing companies using a neural network model. I.V. Arinichev and I.V. Bogdashev, using binary classification trees, built an algorithm for determining the bankruptcy risk of a small business [24]. The study by E. Yu. Makeeva and I.V. Arshavsky [22] focuses on the role of qualitative information in the analysis of the company bankruptcy risk. This information is not directly reflected in the company's financial performance but is present in corporate annual reports. Based on the application of methods of semantic analysis of the company's reporting and an ensemble of artificial neural networks, it was concluded that the predictive ability of the model increases when high-quality information is included in it.

Separately, we note the study of E.A. Fedorova, S.O. Musienko, and F. Yu. Fedorov [19]. This is one of the few papers in which, for Russian small and medium-sized businesses, using machine learning methods, the standard values of indicators appearing in bankruptcy legislation, as well as in foreign studies of the risk of bankruptcy of organizations, are specified. We mainly considered indicators of liquidity and financial stability. The methods of the decision tree, random forest, bagging, and boosting were used. The random forest method showed the best predictive ability.

This study is also devoted to identifying the most significant factors of the bankruptcy of Russian companies and determining their critical values. In contrast to [19], the object of the study is Russian large commercial companies in various industries. A wide range of both financial and non-financial factors of the potential bankruptcy of a company is considered. In accordance with the purpose of the article, the random forest method is used as the main research tool.

RANDOM FOREST METHOD

A random forest is a classification algorithm that consists of many decision trees. This algorithm is used for classification, clustering, feature selection, and anomaly search. The random forest also determines the importance of the factors, i.e. their influence on the classification process. Thus, it becomes possible to arrange the factors in order of priority [25].

Compared to regression models that are relatively sensitive to outliers, random forest is more robust to this problem. Another advantage of the random forest method is its lower susceptibility to overfitting compared to the neural network method [7].

The basic unit of the random forest algorithm is the decision tree. This is a series of questions about the input that can be answered with yes or no. Questions are asked until the tree comes to a decision. Random forest is a machine learning algorithm, and its main advantage is the ability to process new data that has never been seen before. Therefore, so that the model does not overfit, not one decision tree is used, but a random forest consisting of them. *Fig. 1* shows a conditional example of a random forest. Yellow marks the solutions that the trees come to.

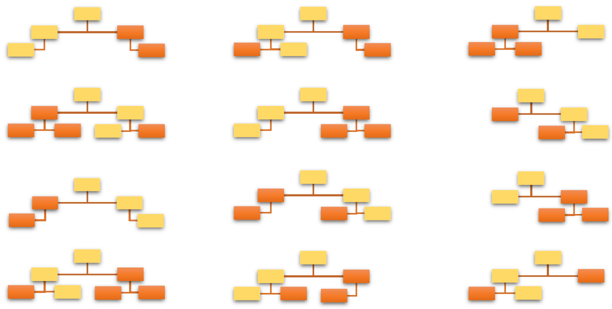

Fig. 1. **Random forest algorithm prototype**

Source: compiled by the authors.

The general principle of the random forest algorithm is that the researcher builds many decision trees (classification trees). Each tree in the random forest returns a prediction of the result, and the result with the most votes becomes the prediction of the forest. The key feature is the low correlation between trees. This effect is due to the fact that different trees answer different "yes" or "no" questions, thereby coming to a result. Some trees may lead to a false result and some to the correct one. Thus, the trees protect each other from individual errors. A large number of weakly correlated trees parsing information together will outperform any of their individual constituents [26].

STAGES OF EMPIRICAL RESEARCH

In accordance with the purpose of the study, the following stages of research were carried out.

At the *first stage*, a preliminary set of quantitative and qualitative variables for analysis was determined. The list of bankruptcy factors of a company included variables from various domestic and foreign studies (see, for example, [1, 6, 27–29]). All variables were divided into five groups: *indicators of profitability, liquidity, turnover, financial stability, and general (non-financial)* indicators. Moreover, in order to avoid the "cannibalization effect", a number of variables were removed, since the analysis of descriptive statistics revealed that they have a high correlation with other variables.

The final composition of variables for the random forest algorithm, taking into account the exclusion of some of the variables, is presented in *Table 1*.

At the *second stage,* using the random forest method, the most significant (important) factors for the sample companies were determined in terms of their impact on the probability of bankruptcy. This was done using the Scikit-learn library, whose algorithms allow you to calculate the relative importance of each factor, i.e. the contribution of each indicator to the prediction. Random Forest uses the Gini index to assess the significance of factors. The obtained values of the estimates of the significance of individual indicators were normalized so that the sum of all estimates was equal to 1. The higher the value of the assessment of the significance of a particular factor, the greater its contribution to the prediction of the company bankruptcy. The significance of features was determined both for the entire sample and for industry subsamples — for manufacturing companies, service companies, and retail companies.

At the *third stage*, the threshold values of indicators were determined, which, according to the results of the analysis of individual industries, turned out to be the most significant for assessing the bankruptcy risk of a company. To do this, we used a visual analysis of a classification model built according to the random forest algorithm using the PDPbox (Python) library. This result allows us to divide companies into three groups according to the value of individual

Table 1

Variables used in the study

Number	Indicator type	Formula	Variable type
1	Profitability	$\frac{EBITDA_{i,p}}{A_{i,p}}$, where $EBITDA_{i,p}$ – earnings before taxes, interest, and depreciation of a company *i* per year *p*; $A_{i,p}$ – the total assets of a company *i* per year *p*	Quantitative
2	Profitability	$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$, where $Retained\ earnings_{i,p}$ – retained earnings of a company *i* per year *p*; $A_{i,p}$ – the total assets of a company *i* per year *p*	Quantitative
3	Profitability	$\frac{EBIT_{i,p}}{Rev_{i,p}}$, where $EBIT_{i,p}$ – earnings before taxes and interest of a company *i* per year *p*; $Rev_{i,p}$ — ototal revenue of a company *i* per year *p*	Quantitative
4	Profitability	$ROE_{i,p} = \frac{NI_{i,p}}{Equity_{i,p}}$, where $ROE_{i,p}$ – return on equity of a company *i* per year *p*; $NI_{i,p}$ — net income of company *i* per year *p*; $Equity_{i,p}$ – equity of a company *i* per year *p*	Quantitative
5	Liquidity	$\frac{CA_{i,p}}{STD_{i,p}}$, where $CA_{i,p}$ – the current assets of a company *i* per year *p*; $STD_{i,p}$ – short-term debt of a company *i* per year *p*	Quantitative
6	Liquidity	$\frac{Cash_{i,p}}{A_{i,p}}$, where $Cash_{i,p}$ – cash of a company *i* per year *p*; $A_{i,p}$ – total assets of a company *i* per year *p*	Quantitative
7	Financial stability	$\frac{FA_{i,p}}{CA_{i,p}}$, where $FA_{i,p}$ – fixed assets of a company *i* per year *p*; $CA_{i,p}$ – current assets of a company *i* per year *p*	Quantitative
8	Financial stability	$\frac{EBITDA_{i,p}}{Int_{i,p}}$, where $EBITDA_{i,p}$ – earnings before taxes, interest, and depreciation of a company *i* per year *p*; $Int_{i,p}$ – interest payable to a company *i* per year *p*	Quantitative

Table 1 (continued)

Number	Indicator type	Formula	Variable type
9	Financial stability	$\frac{WC_{i,p}}{LTD_{i,p}}$, where $WC_{i,p}$ – working capital of a company *i* per year *p*; $LTD_{i,p}$ — long-term debt of a company *i* per year *p*	Quantitative
10	Financial stability	$\frac{STD+LTD_{i,p}}{Equity_{i,p}}$, where $STD+LTD_{i,p}$ – total debt of a company *i* per year *p*; $Equity_{i,p}$ — equity of a company *i* per year *p*	Quantitative
11	Turnover	$\frac{Rev_{i,p}}{\overline{A_{i,p}}}$, where $Rev_{i,p}$ – revenue of a company *i* per year *p*; $\overline{A_{i,p}}$ — average assets of a company *i* per year *p*	Quantitative
12	Turnover	$Inventories\,turnover_{i,p} = \frac{COGS_{i,p}}{Average\,inventories_{i,p}}$, where $COGS_{i,p}$ – cost of a company *i* per year *p*; $Average\,inventories_{i,p}$ – average inventories of a company *i* per year *p*	Quantitative
13	Turnover	$Fixed\,assets\,turnover_{i,p} = \frac{Revenue_{i,p}}{Average\,fixed\,assets_{i,p}}$, where $Revenue_{i,p}$ – revenue of a company *i* per year *p*; $Average\,fixed\,assets_{i,p}$ — average fixed assests of a company *i* per year *p*	Quantitative
14	Turnover	$\begin{matrix} Accounts\,receivable \\ turnover\,period \end{matrix}_{i,p} = \frac{Revenue_{i,p}}{Average\,receivables_{i,p}}$, where $Revenue_{i,p}$ – revenue of a company *i* per year *p*; $Average\,receivables_{i,p}$ — average receivables of a company *i* per year *p*	Quantitative
15	Turnover	$\frac{Rev_{i,p}}{WC_{i,p}}$, where $Rev_{i,p}$ — total revenue of a company *i* per year *p*; $WC_{i,p}$ – working capital of a company *i* per year *p*	Quantitative
16	General	$\ln\left(\frac{A_{i,p}}{GPN\ price\,index_p}\right)$, where $A_{i,p}$ – assets of a company *i* per year *p*; $GPN\ price\ index_p$ – GDP price deflator per year *p*	Quantitative

Table 1 (continued)

Number	Indicator type	Formula	Variable type
17	General	$bin_{1,i,p}$ (takes value 1, if $NI_{i,p} + NI_{i-1,p} < 0$ and value 0 otherwise, where $NI_{i,p}$ – net income of a company *i* per year *p*)	Binary
18	General	$bin_{2,i,p}$ (takes value 1, if $Total\ liabilities_{i,p} > Total\ Assets_{i,p}$ and value 0 otherwise, where $Total\ liabilities_{i,p}$ – current debt of a company *i* per year *p*, $Total\ Assets_{i,p}$ – total assets of a company *i* per year *p*)	Binary
19	General	$\frac{NI_{i,p} - NI_{i-1,p}}{\lvert NI_{i,p}\rvert + \lvert NI_{i-1,p}\rvert}$, where $NI_{i,p}$ – net income of a company *i* per year *p*	Quantitative
20–23	General	Total headcount (4 variables: 251–500 people, 501–1000 people, 1001–5000 people, more than 5000 people)	Binary
24–27	General	Type of business operation (4 variables: manufacturing, retail, services, etc.)	Binary
28–31	General	Organizational and legal form (4 variables: LLC, PJSC, JSC, etc.)	Binary
32–36	General	Type of ownership (5 variables: private, foreign, Russian+foreign, state/federal/ municipal property, etc.)	Binary

Source: compiled by the authors according to [1, 6, 27–29].

indicators — with high, medium, and low risks of bankruptcy.

We note that in this paper, the sample was divided into *test* and *training* subsamples. For this purpose, the train_test_split module of the Scikit-learn library was used. With the help of this module, 20% of the test values and 80% of the training values were extracted from all data. In order for the test values to cover different years, the train_test_split function selects 20% of the data for the test sample *randomly*. This splitting of the sample results in the model being trained on one piece of data and tested on another. This avoids the effect of retraining the model and makes sure that the constructed model not only "remembers the answers" in the training set, but also predicts the results on the test part of the set with high accuracy.

SAMPLE

The sample included large Russian companies. According to the Unified Register of Small and Medium Businesses, medium-sized businesses include companies with up to 250 employees and with annual revenue of up to 2 billion rubles.[4] Thus, in this study, companies with revenues of over 2 billion rubles per year and more than 250 employees were analyzed. The SPARK system was used to collect data for 12 years (from 2009 to 2020).[5] The sample included data on 3184 companies. The total number of observations was 38,208. According to the collected data, 68% of the companies

[4] Unified register of small and medium-sized businesses. URL: https://ofd.nalog.ru/about.html?section=conditions (accessed on 11.03.2021).

[5] SPARK database. URL: http://www. spark-interfax.ru/ (accessed on 15.04.2021).

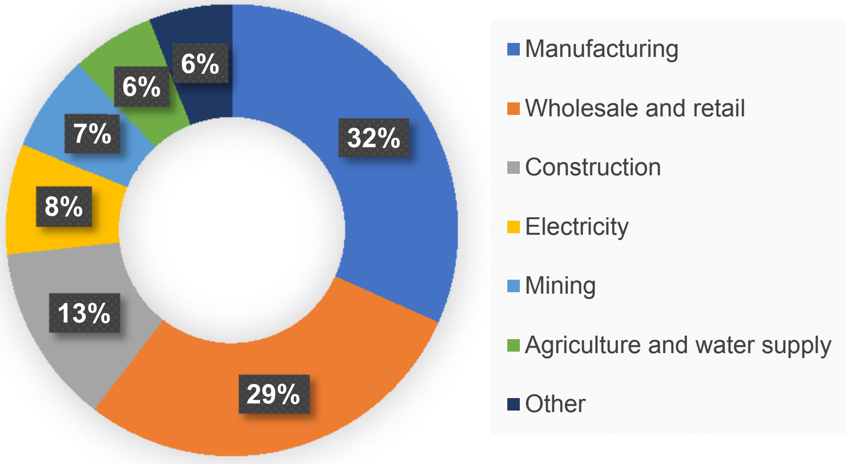

Fig. 2. **Distribution of sample companies by industry**

Source: compiled by the authors.

included in the sample continued operations throughout the entire observation period, and 32% went bankrupt.

The sample of the study included companies in 7 industries: manufacturing, retail, construction, electricity, mining, agriculture, and water supply, as well as other industries, which included companies in education, healthcare, agriculture, hospitality, etc. The diagram in *Fig. 2* shows the distribution of companies by industry.

Fig. 2 shows that the sample contains a high share of companies from the retail sector, as well as the manufacturing industry, which together make up 61% of the entire sample. Construction and electric power companies also account for a significant share.

In order to analyze possible differences in the risk factors for bankruptcy of companies by industry, three large industry groups of companies were identified: manufacturing, retail, and services. Thus, the sample was divided into three approximately equal parts (*Fig. 3*).

The sample mainly includes companies with 1,001 to 5,000 employees — 41%. More than 50% are companies with the legal form of LLC. Most of the companies in the sample are privately owned by citizens of the Russian Federation — 69%. 13% of companies are owned by foreign investors (citizens, states, and legal entities).

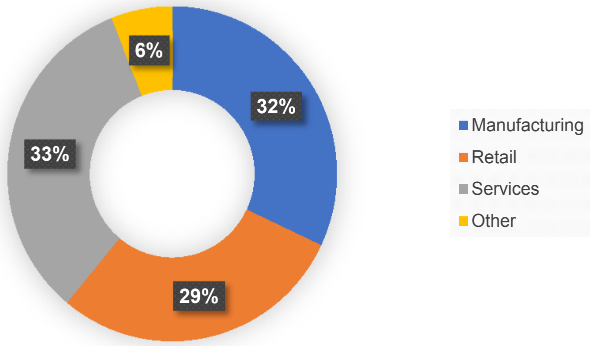

Fig. 3. **Distribution of sample companies by type of business operation**

Source: compiled by the authors.

Descriptive statistics of the studied variables are presented in *Table 2*.

Table 2 indicates that the sample included both profitable and non-profitable companies, i.e. companies with both positive and negative profitability ratios.

RESEARCH RESULTS

The random forest method was used to predict the bankruptcy of companies over the period under review, as well as to determine the factors that most affect the probability of bankruptcy. As for the results, first of all, it should be noted that in 86% of cases the model correctly predicted whether the sample company went bankrupt during the period of observation or whether it continued to

Table 2

Descriptive statistics

Indicator	Type of indicator	Sample average	Standard deviation	Min	Max
$\frac{EBITDA_{i,p}}{A_{i,p}}$	Profitability	0.089	0.145	−1.029	1.213
$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Profitability	0.214	0.361	−2.448	1.073
$\frac{EBIT_{i,p}}{Rev_{i,p}}$	Profitability	0.067	0.237	−3.936	2.635
$ROE_{i,p}$	Profitability	0.161	0.765	−10.211	9.333
$\frac{CA_{i,p}}{STD_{i,p}}$	Liquidity	2.115	2.578	0.067	31.779
$\frac{Cash_{i,p}}{A_{i,p}}$	Liquidity	0.049	0.084	0	0.749
$\frac{FA_{i,p}}{CA_{i,p}}$	Financial stability	1.281	2.234	0	24.831
$\frac{EBITDA_{i,p}}{Int_{i,p}}$	Financial stability	33.742	163.636	−336.7	2536
$\frac{WC_{i,p}}{LTD_{i,p}}$	Financial stability	23.457	124.003	−967.833	1513
$\frac{STD+LTD_{i,p}}{Equity_{i,p}}$	Financial stability	6.283	25.514	−273.306	364.871
$\frac{Rev_{i,p}}{A_{i,p}}$	Turnover	1.860	1.679	0	13.919

Table 2 (continued)

Indicator	Type of indicator	Sample average	Standard deviation	Min	Max
$Inventories\ turnover_{i,p} = \dfrac{COGS_{i,p}}{Average\ inventories_{i,p}}$	Turnover	36.086	148.735	0	1954
$Fixed\ assets\ turnover_{i,p} = \dfrac{Revenue_{i,p}}{Average\ fixed\ assets_{i,p}}$	Turnover	40.239	174.057	0.001	2595
$\begin{pmatrix} Accounts\ receivable \\ turnover\ period \end{pmatrix}_{i,p} = \dfrac{Revenue_{i,p}}{Average\ receivables_{i,p}}$	Turnover	9.020	12.493	0	138.983
$\ln\left(\dfrac{A_{i,p}}{GPN\ price\ index_p}\right)$	General	17.884	1.524	9.158	25.654
$\dfrac{NI_{i,p} - NI_{i-1,p}}{\lvert NI_{i,p}\rvert + \lvert NI_{i-1,p}\rvert}$	General	0.003	0.605	−1	1

Source: compiled by the authors.

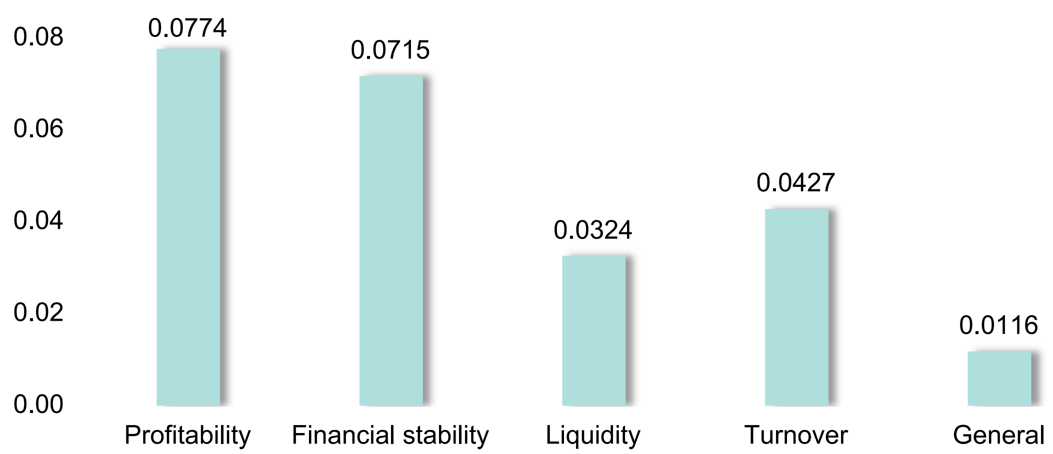

Fig. 4. **Evaluation of the importance of groups of bankruptcy risk factors according to the random forest algorithm**

Source: compiled by the authors based on the research results.

Table 3

Evaluation of the importance of bankruptcy risk factors for the whole sample according to the random forest algorithm

Number	Type of indicator	Designation	Type of variable	Importance
1	Profitability	$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Quantitative	0.235
2	Financial stability	$\frac{STD+LTD_{i,p}}{Equity_{i,p}}$	Quantitative	0.141
3	General	$ln\left(\frac{A_{i,p}}{GPD_{p}}\right)$	Quantitative	0.112
4	Turnover	$\frac{Rev_{i,p}}{WC_{i,p}}$	Quantitative	0.073
5	Financial stability	$\frac{FA_{i,p}}{CA_{i,p}}$	Quantitative	0.052
6	Turnover	$\left(\begin{matrix}Accounts\ receivable\\ turnover\ period\end{matrix}\right)_{i,p} = \frac{Revenue_{i,p}}{Average\ receivables_{i,p}}$	Quantitative	0.051
7	Turnover	$Inventories\ turnover_{i,p} = \frac{COGS_{i,p}}{Average\ inventories_{i,p}}$	Quantitative	0.046
8	Liquidity	$\frac{Cash_{i,p}}{A_{i,p}}$	Quantitative	0.039
9	Profitability	$\frac{EBITDA_{i,p}}{A_{i,p}}$	Quantitative	0.037
10	General	ПАО	Binary	0.034
11	Turnover	$\frac{Rev_{i,p}}{A_{i,p}}$	Quantitative	0.028
12	Profitability	$ROE_{i,p}$	Quantitative	0.027
13	Liquidity	$\frac{CA_{i,p}}{STD_{i,p}}$	Quantitative	0.026

Table 3 (continued)

Number	Type of indicator	Designation	Type of variable	Importance
14	Financial stability	$\frac{WC_{i,p}}{LTD_{i,p}}$	Quantitative	0.022
15	Turnover	$Fixed\ assets\ turnover_{i,p} = \frac{Revenue_{i,p}}{Average\ fixed\ assets_{i,p}}$	Quantitative	0.015

Source: compiled by the authors based on the research results.

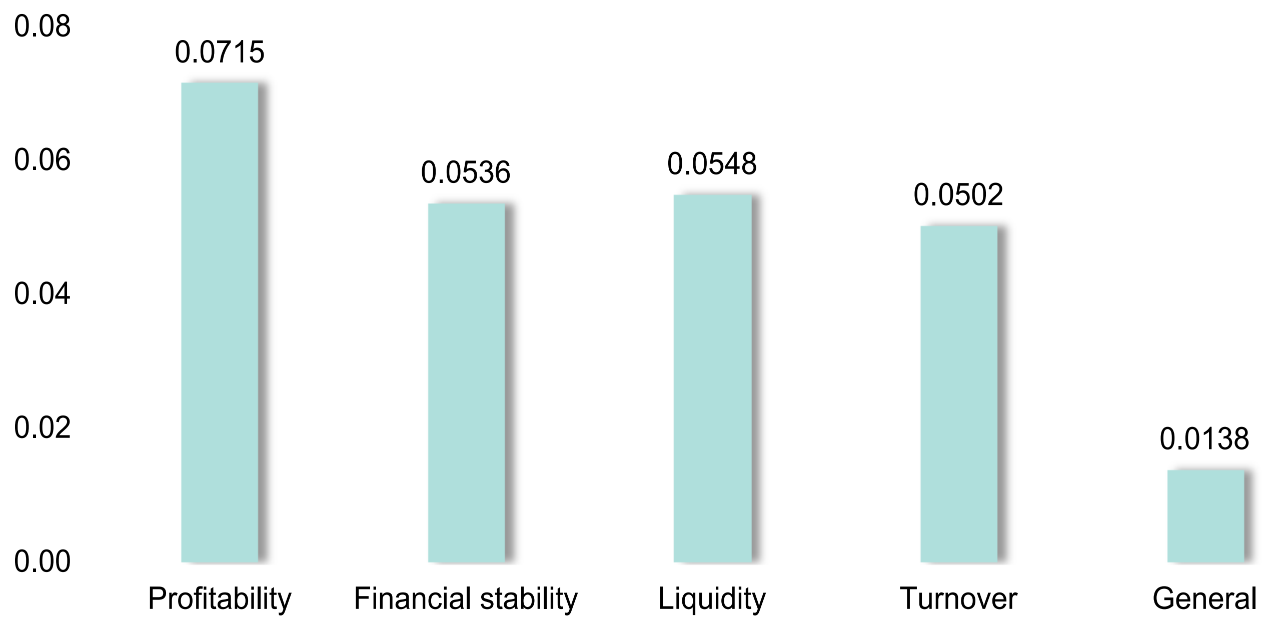

Fig. 5. **Evaluation of the importance of groups of bankruptcy risk factors according to the random forest algorithm for manufacturing companies**

Source: compiled by the authors based on the research results.

operate. This percentage can be considered quite high since in a number of domestic studies the prediction accuracy of machine learning methods varies from 73 to 90% [8, 31].

Further, using the Gini index, the significance of each of the five considered groups of indicators for predicting bankruptcy for the entire sample was assessed. *Fig. 4* presents a graph showing the average value of the indicator of each category. Since each group had a different number of indicators, to evaluate the importance of the indicators of each group, the average influence of the factors of each group on the probability of bankruptcy was calculated. To do this, the overall evaluation of the importance of factors belonging to a particular group was divided by the number of factors of this group in the model.

Fig. 4 shows that the most significant categories of indicators in terms of predicting the probability of bankruptcy of Russian

Table 4

Evaluation of the importance of bankruptcy risk factors for the whole sample according to the random forest algorithm for manufacturing companies

Number	Type of indicator	Designation	Type of variable	Importance
1	Profitability	$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Quantitative	0.182
2	Liquidity	$\frac{Cash_{i,p}}{A_{i,p}}$	Quantitative	0.089
3	Financial stability	$\frac{STD+LTD_{i,p}}{Equity_{i,p}}$	Quantitative	0.088
4	Turnover	$Inventories\ turnover_{i,p} = \frac{COGS_{i,p}}{Average\ inventories_{i,p}}$	Quantitative	0.082
5	Turnover	$\left(\begin{matrix}Accounts\ receivable\\ turnover\ period\end{matrix}\right)_{i,p} = \frac{Revenue_{i,p}}{Average\ receivables_{i,p}}$	Quantitative	0.067
6	General	$ln\left(\frac{A_{i,p}}{GPD_{p}}\right)$	Quantitative	0.061
7	Turnover	$\frac{Rev_{i,p}}{A_{i,p}}$	Quantitative	0.047
8	Financial stability	$\frac{FA_{i,p}}{CA_{i,p}}$	Quantitative	0.045
9	Profitability	$\frac{EBIT_{i,p}}{Rev_{i,p}}$	Quantitative	0.042
10	Profitability	$ROE_{i,p}$	Quantitative	0.032
11	Profitability	$\frac{EBITDA_{i,p}}{A_{i,p}}$	Quantitative	0.03
12	Turnover	$Fixed\ assets\ turnover_{i,p} = \frac{Revenue_{i,p}}{Average\ fixed\ assets_{i,p}}$	Quantitative	0.029

Table 4 (continued)

Number	Type of indicator	Designation	Type of variable	Importance
13	Financial stability	$\frac{WC_{i,p}}{LTD_{i,p}}$	Quantitative	0.028
14	Turnover	$\frac{Rev_{i,p}}{WC_{i,p}}$	Quantitative	0.025
15	General	ПАО	Binary	0.023
16	General	Russian + foreign ownership	Binary	0.02
17	Liquidity	$\frac{CA_{i,p}}{STD_{i,p}}$	Quantitative	0.019
18	General	Headcount 501–1000	Binary	0.019
19	General	Headcount 251–500	Binary	0.014
20	General	АО	Binary	0.01

Source: compiled by the authors based on the research results.

companies are indicators of profitability and financial stability. The evaluation of their importance is almost twice as high as the evaluation of the importance of liquidity and turnover indicators.

The relative influence of individual indicators on the probability of bankruptcy is given in *Table 3*. This and subsequent tables present the most significant factors, the total importance of which according to the Gini index is 0.9.

Table 3 shows that the total weight of the indicators included in the profitability and financial stability groups is almost 53%, which confirms the need to take these indicators into account when predicting bankruptcy. The relative influence of one indicator in these groups is 8 and 7%, respectively. It should be emphasized that in the classical Altman model, most of the indicators are related to profitability and financial stability [1]. Our result is also consistent with the study by A. M. Karminsky and R. N. Burekhin, who showed that indicators of financial stability and liquidity are the most significant for predicting the bankruptcy of a company in the Russian market [7]. B. B. Demeshev in 2014, analyzing the construction industry in Russia, also emphasized the importance of the profitability indicator, which in his work was calculated as the ratio of earnings before interest and taxes to the total assets of the company [20].

It is also interesting that “general” indicators (for example, the number of employees, form of ownership, etc.) turned out to be insignificant when predicting the probability of bankruptcy. It should be noted that non-financial indicators turned out to be insignificant in some other studies of Russian companies. This, in particular, suggests that “the large size and long period of operation of the company in the market cannot guarantee stability in the Russian market” [7].

The results for the subsample of manufacturing companies are similar to the

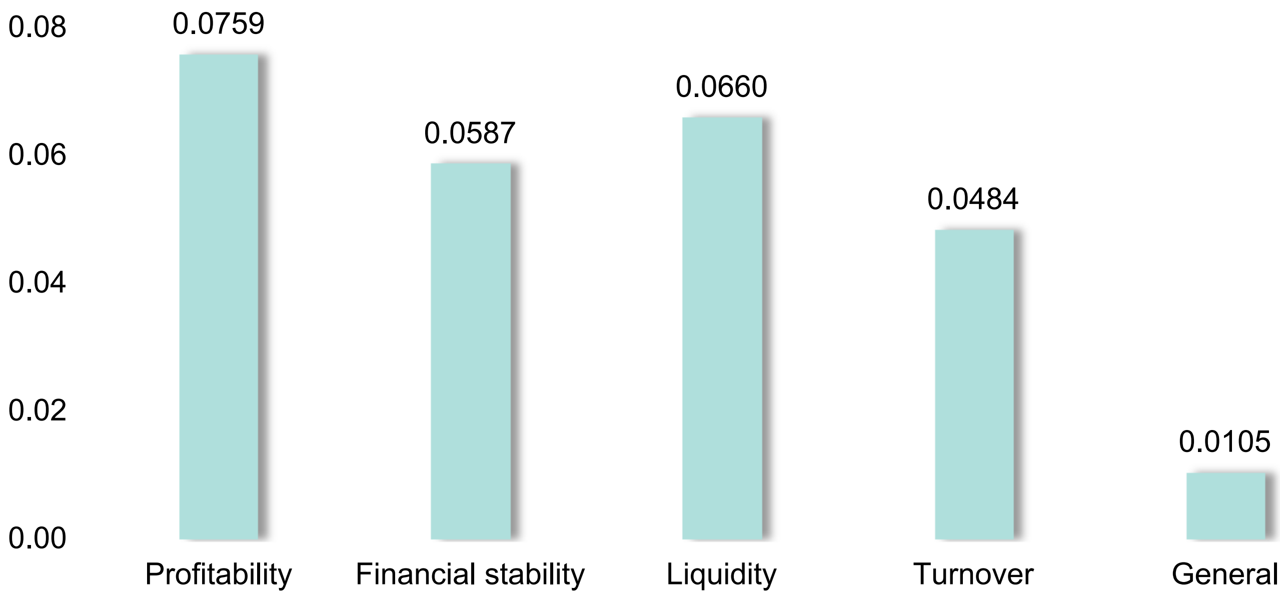

Fig. 6. **Evaluation of the importance of groups of bankruptcy risk factors according to the random forest algorithm for retail companies**

Source: compiled by the authors based on the research results.

Table 5

Evaluation of the importance of bankruptcy risk factors according to the random forest algorithm for retail companies

Number	Type of indicator	Designation	Type of variable	Importance
1	Profitability	$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Quantitative	0.219
2	Financial stability	$\frac{WC_{i,p}}{LTD_{i,p}}$	Quantitative	0.082
3	Liquidity	$\frac{Cash_{i,p}}{A_{i,p}}$	Quantitative	0.079
4	General	$\ln\left(\frac{A_{i,p}}{GPD_p}\right)$	Quantitative	0.075
5	Turnover	$\left(\begin{matrix} Accounts\ receivable \\ turnover\ period \end{matrix}\right)_{i,p} = \frac{Revenue_{i,p}}{Average\ receivables_{i,p}}$	Quantitative	0.064
6	Financial stability	$\frac{FA_{i,p}}{CA_{i,p}}$	Quantitative	0.062

Table 5 (continued)

Number	Type of indicator	Designation	Type of variable	Importance
7	Turnover	$Inventories\,turnover_{i,p} = \frac{COGS_{i,p}}{Average\,inventories_{i,p}}$	Quantitative	0.059
8	Liquidity	$\frac{CA_{i,p}}{STD_{i,p}}$	Quantitative	0.053
9	Turnover	$\frac{Rev_{i,p}}{\overline{A}_{i,p}}$	Quantitative	0.05
10	Turnover	$\frac{Rev_{i,p}}{WC_{i,p}}$	Quantitative	0.048
11	Financial stability	$\frac{STD+LTD_{i,p}}{Equity_{i,p}}$	Quantitative	0.032
12	Profitability	$\frac{EBITDA_{i,p}}{A_{i,p}}$	Quantitative	0.029
13	Profitability	$ROE_{i,p}$	Quantitative	0.028
14	Profitability	$\frac{EBIT_{i,p}}{Rev_{i,p}}$	Quantitative	0.027

Source: compiled by the authors based on the research results.

results for the entire sample (*Fig. 5*). Again, the most significant groups of indicators in terms of predicting bankruptcy are the groups of profitability and financial stability. However, the role of liquidity and turnover indicators is also significant.

Table 4 shows the importance of individual indicators in predicting the bankruptcy of manufacturing companies.

Table 4 shows that the ratio of retained earnings to assets and financial leverage is still among the most significant indicators. Foreign researchers also emphasize the importance of the indicator of the ratio of retained earnings to assets, since retained earnings demonstrate the size of funds remaining after settlements with all capital providers [31].

As for retail companies, among the considered groups of indicators, profitability ranks first in importance in predicting bankruptcy (*Fig. 6*). The value of liquidity indicators is enhanced in comparison with manufacturing companies, which generally corresponds to the specifics of retail companies.

Table 5 presents the importance of individual indicators in predicting the bankruptcy of retail companies.

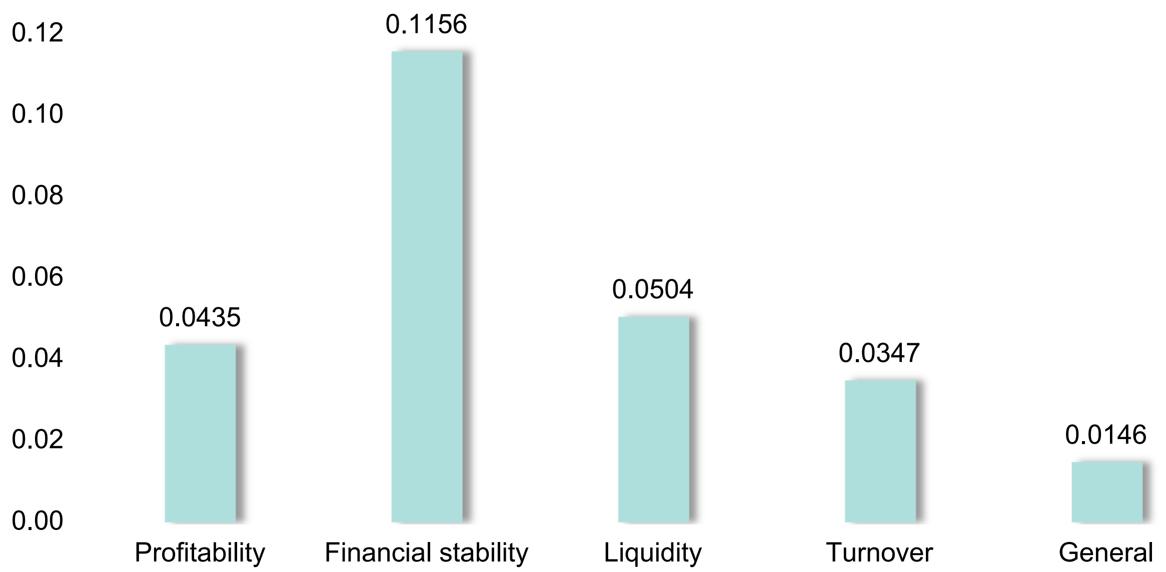

Fig. 7. **Evaluation of the importance of groups of bankruptcy risk factors according to the random forest algorithm for service companies**

Source: compiled by the authors based on the research results.

Table 6

Evaluation of the importance of bankruptcy risk factors according to the random forest algorithm for service companies

Number	Type of indicator	Designation	Type of variable	Importance
1	Financial stability	$\frac{FA_{i,p}}{CA_{i,p}}$	Quantitative	0.165
2	Financial stability	$\frac{STD+LTD_{i,p}}{Equity_{i,p}}$	Quantitative	0.133
3	General	$ln\left(\frac{A_{i,p}}{GPD_{p}}\right)$	Quantitative	0.087
4	Profitability	$\frac{EBIT_{i,p}}{Rev_{i,p}}$	Quantitative	0.077
5	General	ПАО	Binary	0.059
6	Turnover	$Inventories\,turnover_{i,p} = \\ = \frac{COGS_{i,p}}{Average\,inventories_{i,p}}$	Quantitative	0.058
7	Liquidity	$\frac{Cash_{i,p}}{A_{i,p}}$	Quantitative	0.052

Table 6 (continued)

Number	Type of indicator	Designation	Type of variable	Importance
8	Financial stability	$\frac{WC_{i,p}}{LTD_{i,p}}$	Quantitative	0.049
9	Liquidity	$\frac{CA_{i,p}}{STD_{i,p}}$	Quantitative	0.048
10	Profitability	$\frac{EBITDA_{i,p}}{A_{i,p}}$	Quantitative	0.041
11	Turnover	$\left(\begin{array}{c}Accounts\ receivable\\ turnover\ period\end{array}\right)_{i,p} = \frac{Revenue_{i,p}}{Average\ receivables_{i,p}}$	Quantitative	0.039
12	Profitability	$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Quantitative	0.029
13	Profitability	$ROE_{i,p}$	Quantitative	0.028
14	Turnover	$\frac{Rev_{i,p}}{\overline{A_{i,p}}}$	Quantitative	0.027

Source: compiled by the authors based on the research results.

The results of *Table 5* indicate that the first three places in the list of the most important indicators in terms of predicting bankruptcies are occupied by profitability (0.218), financial stability (0.082), and liquidity (0.079), respectively.

In the service sector, among the groups of indicators under review, the financial stability ratios are most significant in predicting bankruptcy, with a large margin from other groups of indicators (*Fig. 7*).

Table 6 shows the importance of individual indicators in predicting the bankruptcy of service companies.

According to *Table 6*, for companies in the service sector, the indicator of the ratio of fixed assets to current assets (0.165) has the greatest significance in predicting bankruptcy. The indicator of the ratio of debt and equity financing also has a high significance (0.133). The indicator of financial leverage has a significant impact on the probability of bankruptcy, with an increase in the level of debt, the company risks reducing its financial stability [32].

In addition to the presented results, threshold values of financial indicators were obtained, delimiting intervals characterized by different degrees of bankruptcy risk. To do this, we used a visual analysis of a classification model built according to the random forest algorithm using the PDPbox (Python) library. The library allows you to build partial dependence plots or PDP, which reflect the assessment of the influence of individual variables on the classification result. The PDP

Table 7

Intervals of indicator values characterized by different degrees of company bankruptcy risk

Indicator	Interval of indicator values		
	With high bankruptcy risk	With medium bankruptcy risk	With low bankruptcy risk
Entire sample			
$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Below 0.087	From 0.087 to 0.156	Above 0.156
$\frac{STD+LTD_{i,p}}{Equity_{i,p}}$	Below 2.4 Above 10.9	From 2.4 to 10.9	–
$\frac{Rev_{i,p}}{WC_{i,p}}$	Below 0.2	From 0.2 to 1.8	Above 1.8
$\frac{Cash_{i,p}}{A_{i,p}}$	Below 0.003	From 0.003 to 0.02	Above 0.02
Manufacturing companies			
$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Below 0.04	From 0.04 to 0.28	Above 0.28
$\frac{Cash_{i,p}}{A_{i,p}}$	Below 0.002	From 0.002 to 0.005	Above 0.005
$\frac{STD+LTD_{i,p}}{Equity_{i,p}}$	Below 2.5 Above 8	From 2.5 to 8	–
$\left(\begin{matrix}Accounts\ receivable\\ turnover\ period\end{matrix}\right)_{i,p} = \frac{Revenue_{i,p}}{Average\ receivables_{i,p}}$	Below 1.89	From 1.89 to 4.73	Above 4.73
Retail companies			
$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Below 0.12	From 0.12 to 0.27	Above 0.27
$\frac{WC_{i,p}}{LTD_{i,p}}$	Below –0.5	From –0.5 to 1.6	Above 1.6

Table 7 (continued)

Indicator	Interval of indicator values		
	With high bankruptcy risk	With medium bankruptcy risk	With low bankruptcy risk
$\frac{Cash_{i,p}}{A_{i,p}}$	Below 0.0004	From 0.0004 to 0.004	Above 0.004
$Accounts\ receivable\ turnover\ period_{i,p} = \frac{Revenue_{i,p}}{Average\ receivables_{i,p}}$	Below 0.13	From 0.13 to 2.72	Above 2.72
Service companies			
$\frac{Fixed\ assets_{i,p}}{Current\ assets_{i,p}}$	Below 0.01	From 0.01 to 0.73	Above 0.73
$\frac{EBIT_{i,p}}{Rev_{i,p}}$	Below 0.134	From 0.134 to 0.323	Above 0.323
$Inventories\ turnover_{i,p} = \frac{COGS_{i,p}}{Average\ inventories_{i,p}}$	Below 3.4	From 3.4 to 11.8	Above 11.8
$\frac{Cash_{i,p}}{A_{i,p}}$	Below 0.001	From 0.001 to 0.01	Above 0.01

Source: compiled by the authors based on the research results.

is a broken line. Although the random forest model does not indicate the direction of the relationship of features with the classification result, this can be seen on the PDP plots using signs similar to the signs in front of the coefficients in regression models.

When analyzing financial insolvency, PDPbox helps to trace the relationship between one bankruptcy factor in the classification model and the possible bankruptcy of the company. The set of admissible values of the attribute is divided into three areas — values characterized by "high", "medium" and "low" bankruptcy risk, respectively. The areas correspond to qualitatively different groups of financial condition. For division into areas, an empirical method (elbow method) is used, which allows you to find "critical" points on the RDP. These points (as well as the regions themselves) are listed in the table for each object separately.

Visually, on the PDP plots of the dependence of the probability of bankruptcy on the value of the factor, one can observe a noticeable change in the "curvature". Starting from a certain value of the factor, the probability of bankruptcy significantly decreases or increases. Such a change in "curvature" is analyzed using the values of the cosines of the angles between the links of the PDP. To determine the threshold values that form the boundaries of the intervals that define qualitatively different groups, two nodes with the largest cosine of the angle are selected.

Table 7 shows the intervals of the values of the indicators of each of the groups that turned out to be the most important for predicting bankruptcy, depending on the degree of bankruptcy risk.

The results presented in *Table 7* allow getting an idea of what intervals of values of financial indicators are characterized by different risks of bankruptcy for the company. This result allows us to classify companies depending on the degree of probability of their bankruptcy. To obtain more detailed conclusions, it is advisable to conduct a similar analysis for individual sectors/sub-sectors of the Russian economy.

CONCLUSIONS

Assessing the probability of a company going bankrupt and identifying the bankruptcy risk factors are extremely important for understanding business prospects in any industry. Predicting the probability of bankruptcy of a company is traditionally done using mathematical models, usually based on econometric methods or machine learning methods. These models are adapted to the specifics of the markets of individual countries, which involves testing the model on data collected from a sample of companies in the respective country.

Recent research shows that machine learning methods provide higher accuracy in predicting the probability of a company going bankrupt compared to econometric methods. At the same time, in most studies on Russian companies based on machine learning methods, the predictive ability of various models is compared, while relatively little attention is paid to the analysis of individual factors of company bankruptcy.

As part of this study, we solved the problem of assessing the relative influence of individual factors on the probability of bankruptcy of large Russian companies in various industries using one of the machine learning methods, the random forest algorithm. The indicators for 3184 companies from 2009 to 2020 were considered. *For the entire sample*, the tested model correctly predicted the possibility of a company going bankrupt in 86% of cases over the period under review. This result confirmed that machine learning methods are *highly effective* (and, in particular, the random forest algorithm) in solving the problem of bankruptcy prediction for a company.

The study also showed that the bankruptcy risk factors of companies significantly depend on their industry affiliation:

For *manufacturing companies*, first of all, attention should be paid to a group of profitability indicators. For example, according to the results, the bankruptcy risk increases significantly when the ratio of retained earnings to assets is less than 4%.

For *retail companies*, it is necessary, first of all, to focus on liquidity and profitability indicators. The ratio of cash to assets must not be less than 0.04%, and the ratio of retained earnings to assets must not be less than 12%.

For *service companies*, indicators of financial stability are a priority in terms of predicting the probability of bankruptcy. In particular, a low risk of bankruptcy is observed when the ratio of fixed assets to current assets exceeds 73%.

Thus, the paper not only identifies factors that significantly affect the probability of bankruptcy of Russian companies in various industries but also determines the "threshold" values of these indicators, at which the risk of bankruptcy increases significantly.

The research results can be used by both internal (management, board of directors) and external (analysts, creditors, etc.) stakeholders to determine the current financial condition of the company, as well as forecast business development prospects.

Possible directions for further research include an in-depth analysis of the factors of bankruptcy of small and medium-sized Russian enterprises. Also, in order to obtain more detailed conclusions, it is advisable to analyze the bankruptcy factors for individual sectors/sub-sectors of the Russian economy with machine learning methods. The solution and identification of the bankruptcy risk factors of Russian companies should lead to a reduction in the number of bankrupt enterprises, which, in turn, will contribute to the recovery and development of the national economy.

REFERENCES

1. Altman E.I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*. 1968;23(4):589–609. DOI: 10.1111/J.1540–6261.1968.TB 00843.X
2. Altman E.I., Fargher N., Kalotay E. A simple empirical model of equity-implied probabilities of default. *The Journal of Fixed Income*. 2011;20(3):71–85. DOI: 10.3905/jfi.2011.20.3.071
3. Altman E.I., Iwanicz-Drozdowska M., Laitinen E., Suvas A. Distressed firm and bankruptcy prediction in an international context: A review and empirical analysis of Altman's Z-score model. *SSRN Electronic Journal*. 2014. DOI: 10.2139/ssrn.2536340
4. Agarwal V., Taffler R.J. Twenty-five years of the Taffler Z-score model: Does it really have predictive ability? *Accounting and Business Research*. 2007;37(4):285–300. DOI: 10.1080/00014788.2007.9663313
5. Zmijewski M. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research.* 1984;22:59–82. DOI: 10.2307/2490859
6. Ohlson J. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*. 1980;18(1):109–131. DOI: 10.2307/2490395
7. Karminsky A.M., Burekhin R.N. Comparative analysis of methods for forecasting bankruptcies of Russian construction companies. *Business Informatics*. 2019;13(3):52–66. DOI: 10.17323/1998–0663.2019.3.52.66 (In Russ.: *Biznes-informatika.* 2019;13(3):52–66. DOI: 10.17323/1998–0663.2019.3.52.66).
8. Jones S., Hensher D.A. Predicting firm financial distress: A mixed logit model. *The Accounting Review*. 2004;79(4):1011–1038. DOI: 10.2308/accr.2004.79.4.1011
9. Haydarshina G.A. Improvement of methods for assessing risk of bankruptcy of Russian enterprises in modern conditions. *Imushchestvennye otnosheniya v Rossiiskoi Federatsii = Property Relations in the Russian Federation*. 2009;(8):86–95. (In Russ.).
10. Karminsky A.M., Kostrov A.V., Murzenkov T.N. Modeling of the probability of default of Russian banks using econometric methods. Preprint WP7/2012/04. Moscow: NRU HSE; 2012. 64 p. (In Russ.).
11. Behr A., Weinblat J. Default patterns in seven EU countries: A random forest approach. *International Journal of the Economics of Business*. 2017;24(2):181–222. DOI: 10.1080/13571516.2016.1252532
12. Li Y., Wang Y. Machine learning methods of bankruptcy prediction using accounting ratios. *Open Journal of Business and Management*. 2018;6(1):1–20. DOI: 10.4236/ojbm.2018.61001
13. Joshi S., Ramesh R., Tahsildar S. A bankruptcy prediction model using random forest. In: 2nd Int. conf. on intelligent computing and control systems (ICICCS). (Madurai, 14–15 June 2018). Piscataway, NJ: IEEE; 2018. DOI: 10.1109/ICCONS.2018.8663128
14. Denisov D.V., Smirnova D.K. Application of random forest method to estimate the incurred but not reported claims reserve of an insurance company. *International Journal of Open Information Technologies*. 2016;4(7):45–50. (In Russ.).
15. Gruzdev A.V. Random forest method in scoring. *Risk-menedzhment v kreditnoi organizatsii*. 2014;(1):28–43. (In Russ.).
16. Kazakov A.V., Kolyshkin A.V. The development of bankruptcy prediction models in modern Russian economy. *Vestnik Sankt-Peterburgskogo universiteta. Ekonomika = St. Petersburg University Journal of Economic Studies (SUJES)*. 2018;34(2):241–266. (In Russ.). DOI: 10.21638/11701/spbu05.2018.203

17. Kolyshkin A.V., Gilenko E.V., Dovzhenko S.E., Zhilkin S.A., Choe S.E. Forecasting the financial insolvency of enterprises. *Vestnik Sankt-Peterburgskogo universiteta. Ekonomika = St. Petersburg University Journal of Economic Studies (SUJES)*. 2014;(2):122–142. (In Russ.).
18. Fedorova E.A., Gilenko E.V., Dovzhenko S.E. Models for bankruptcy forecasting: Case study of Russian enterprises. *Studies on Russian Economic Development*. 2013;24(2):159–164. (Russ. ed.: *Problemy prognozirovaniya*. 2013;(2):85–92.).
19. Fedorova E.A., Musienko S.O., Fedorov F. Yu. Prediction of bankruptcy of small and medium-sized business entities in Russia. *Finansy i kredit = Finance and Credit*. 2018;24(11):2537–2552. (In Russ.). DOI: 10.24891/fc.24.11.2537
20. Demeshev B., Tikhonova A. Default prediction for Russian companies: Intersectoral comparison. *Ekonomicheskii zhurnal Vysshei shkoly ekonomiki = The HSE Economic Journal*. 2014;18(3):359–386. (In Russ.).
21. Gorbatkov S., Beloliptsev I. A hybrid method for estimating the risk of bankruptcies based on Bayesian neural network ensemble and the logit-model. *Naukovedenie*. 2013;(6). URL: http://naukovedenie.ru/PDF/25EVN 613.pdf (In Russ.).
22. Makeeva E. Yu., Arshavsky I.V. Integration of neural networks and semantic interpretation for bankruptcy prediction. *Korporativnye finansy = Journal of Corporate Finance Research*. 2014;8(4):130–141. (In Russ.). DOI: 10.17323/j.jcfr.2073–0438.8.4.2014.130–141
23. Bogdanova T., Shevgunov T., Uvarova O. Using neural networks for solvency prediction for Russian companies of manufacturing industries. *Business Informatics*. 2013;(2):40–48. (In Russ.: *Biznes-informatika*. 2013;(2):40–48.).
24. Arinichev I.V., Bogdashev I.V. Estimation of bankruptcy risk of small business companies using methods of machine learning. *Vestnik Rossiiskogo universiteta druzhby narodov. Seriya: Ekonomika = RUDN Journal of Economics*. 2017;25(2):242–254. (In Russ.). DOI: 10.22363/2313–2329–2017–25–2–242–254
25. Liaw A., Wiener M. Classification and regression by randomForest. *R News*. 2002;2(3):18–22. URL: https://cogns.northwestern.edu/cbmg/LiawAndWiener2002.pdf
26. Gepp A., Kumar K. Predicting financial distress: A comparison of survival analysis and decision tree techniques. *Procedia Computer Science*. 2015;54:396–404. DOI: 10.1016/j.procs.2015.06.046
27. Altman E.I., Sabato G. Modelling credit risk from SMEs: Evidence from the US market. *ABACUS: A Journal of Accounting, Finance and Business Studies*. 2007;43(3):332–357. DOI: 10.1111/j.1467–6281.2007.00234.x
28. Zhdanov V. Yu., Afanaseva O.A. Bankruptcy risk diagnostics model for aviation enterprises. *Korporativnye finansy = Journal of Corporate Finance Research*. 2011;5(4):77–89. (In Russ.). DOI: 10.17323/j.jcfr.2073–0438.5.4.2011.77–89
29. Drezner Z., Marcoulides G., Stohs M.H. Financial applications of a Tabu search variable selection model. *Journal of Applied Mathematics and Decision Sciences*. 2001;5(4):215–234. DOI: 10.1155/S 1173912601000165
30. Altman E.I., Sabato G., Wilson N. The value of non-financial information in SME risk management. *Journal of Credit Risk*. 2010;6(2):95–127. DOI: 10.21314/JCR.2010.110
31. Pererva O.L., Stepanov S.E., Nezimova S.S. Comparison of econometric models and methods of business analytics for prediction of bankruptcy of enterprises. *Naukovedenie*. 2017;9(6):1–9. URL: https://naukovedenie.ru/PDF/82EVN 617.pdf (In Russ.).
32. Molina C.A. Are firms underleveraged? An examination of the effect of leverage on default probabilities. *The Journal of Finance*. 2005;60(3):1427–1459. DOI: 10.1111/j.1540–6261.2005.00766.x

ABOUT THE AUTHORS

Andrei A. Zhukov — Master program student, Graduate School of Management, St. Petersburg State University, St. Petersburg, Russia
https://orcid.org/0000-0003-1932-0242
Corresponding author
andrey.zhukov399@gmail.com

Egor D. Nikulin — Cand. Sci. (Econ.), Assoc. Prof., Graduate School of Management, St. Petersburg State University, St. Petersburg, Russia
https://orcid.org/0000-0003-0475-3424
nikulin@gsom.spbu.ru

Danil A. Shchuchkin — Master program student, Bonch-Bruevich St. Petersburg State University of Telecommunications, St. Petersburg, Russia
https://orcid.org/0000-0003-3346-1316
da.shchuchkin@gmail.com

Authors' declared contribution:
A.A. Zhukov — developed methodological basis, data collection and analysis, literature review, description of the results.
E.D. Nikulin — wrote the abstract and introduction, literature review, general conclusions and recommendations, methodological basis.
D.A. Shchuchkin — modeling using machine learning methods.

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 25.12.2021; revised on 10.01.2022 and accepted for publication on 27.05.2022. The authors read and approved the final version of the manuscript.