The comparison of empirical methods for modeling credit ratings of industrial companies from BRICS countries

Alexander M. Karminsky, Sergei Grishunin, Natalya Dyachkova & Maxim Bisenov

Eurasian Economic Review
A Journal in Applied Macroeconomics and Finance

ISSN 1309-422X

Eurasian Econ Rev
DOI 10.1007/s40822-019-00130-4

Springer

Eurasian Economic Review
https://doi.org/10.1007/s40822-019-00130-4

ORIGINAL PAPER

The comparison of empirical methods for modeling credit ratings of industrial companies from BRICS countries

Alexander M. Karminsky[1] · Sergei Grishunin[2] · Natalya Dyachkova[1] · Maxim Bisenov[3]

Received: 26 August 2018 / Revised: 15 May 2019 / Accepted: 25 May 2019

Abstract

We compared the ability of various empirical methods to reproduce public credit ratings (PCRs) of industrial companies (ICs) from BRICS countries using publicly available information. This task is important for researchers and practitioners because many of BRICS' ICs lack PCRs from reputable rating agencies such as Moody's, Standard and Poor's, and Fitch. This paper aimed at filling the gap in the existing research as insufficient efforts were focused on prediction of PCRs of ICs from the entire BRICS IC community. The modeled variables are credit ratings (CRs) of 208 BRICS' ICs assigned by Moody's at the year-end from 2006 to 2016. The sample included 1217 observations. Financial explanatory variables included companies' revenue, operating profitability, interest coverage ratio, debt/book capitalization, and cash flow debt coverage. Non-financial explanatory variables included dummies for home region, industry, affiliation with the state, and a set of macroeconomic data of IC's home countries. The set of statistical methods included linear discriminant analysis (LDA), ordered logit regression (OLR), support vector machine (SVM), artificial neural network (ANN), and random forest (RF). The resulting models were checked for in-sample and out-of-sample predictive fit. Our findings revealed that among considered methods of artificial intelligence models (AI), SVM, ANN, and RF outperformed LDA and OLR by predictive power. On testing sample, AI gave on average 55% of precise results and up to 99% with an error within one rating grade; RF demonstrated the best outcome (58% and 100%). Conversely, LDA and OLR on average gave only 37% of precise results and up to 70% with an error within one grade. LDA and OLR also gave higher share of Type I errors (overestimation of ratings) than that of AI. Therefore, AI should have higher practical application than DA and OLR for predicting the ratings of BRICS ICs.

Keywords Credit rating modeling · Industrial company · Linear discriminant analysis · Ordered logit model · Artificial intelligence methods

✉ Natalya Dyachkova
Nfdyachkova@gmail.com

Extended author information available on the last page of the article

Published online: 17 June 2019

JEL Classification C50–C53 · G31–G33

1 Introduction

Higher than average economic growth, strengthening macroeconomic environment in BRICS countries, coupled with increased investor interest, have led to an accelerated trend in growing investments in debt issued by industrial companies (ICs) from these countries. Simultaneously, these assets are the source of the credit risk of significant magnitude. The latter is explained by the various inefficiencies and structural problems in BRICS economies and capital markets (Staples et al. 2013). To minimize credit losses, the debtholders of BRICS ICs badly need reliable tools to assess and forecast the creditworthiness of these assets.

One of the tools that fit the above-mentioned purposes is the public credit ratings (PCRs) assigned by "big-3" international credit rating agencies (ICRAs): Moody's, Standard and Poor's (S&P) or Fitch Ratings (Fitch). These ratings determine the grades to which the debt instruments belong, based on their probability of default (Karminsky and Polozov 2016). The PCRs, assigned by ICRAs, proved their ability to effectively discriminate between defaulters and non-defaulters (including ICs from BRICS) while reflecting more permanent changes in credit risks (Karminsky and Polozov 2016).

However, the large number of ICs' debt in BRICS remained uncovered with CRs from ICRAs. This is underpinned by (1) the significant direct and indirect cost of the rating process for the issuers; and/or (2) the restrictions on the operations of international CRAs in some BRICS countries such as Russia. In absence of PCRs, the debtholders, to assess creditworthiness of the asset, must construct internal credit ratings (ICRs) replicating the missed PCRs. The ICRs are easy to use, low cost and require limited involvement of experts. Having a stable methodology from reputable CRA as a base for the modeling also helps to quick replication of the model.

The well-proven method of modeling of reliable ICRs is the reproduction of missed PCRs with various econometric models from publicly available information (issuers' financial statements, macroeconomic and industry data, etc.) (Karminsky and Peresetsky 2007; Karminsky and Khromova 2016). In these settings, ICRs constitute the forecast of a relative creditworthiness of the debt instrument in the next 12–24 month expressed by the symbol system. The debtholders, knowing the ICR level, can infer the probability of default of the asset from the statistics published by ICRAs (see "Appendix").

However, our research shows that the debtholders of ICs from BRICS face the following problem: what econometric models to apply to ensure that ICRs accurately reflect the creditworthiness of the assets as if they were assessed with the PCRs? The goals of this paper, though, are (1) to compare the ability of various statistical methods to replicate Moody's PCRs for debt issued by BRICS-based ICs and (2) to select the statistical methods, which produces ICRs with the highest predictive power. The relevance of the paper is determined by (1) the solution of one of the critical problems of ICR modeling and (2) the narrow research in this area (see discussion below).

The novelty of the paper is the application and comparison of wide range of statistical models for ICR forecasting while the peers applied only narrow set of models (mainly OLR or OPR). It is also driven by (1) the selection of explanatory variables in the models that provide the best match to the credit factors listed in Moody's rating methodologies; (2) the study of diversified sample of ICs from all BRICS countries (218 ICs from 12 industries); and (3) the application of the most actual data (for the period from 2006 to 2016) to control for the data stationarity.

The rest of the paper is structured as follows. Section 2 represents the literature review in field of ICR modeling. Section 3 explores the data, the set of explanatory variables and the methods of modeling. Finally, in Sect. 4, the accuracy of each method of reproducing Moody's PCRs is discussed and conclusions are formulated.

2 Modeling the credit ratings of international credit rating agencies: the literature review

Majority of the efforts point out the growing magnitude of credit risks in business and importance of its modeling from publicly available information (Jarrow 2009). The substantial number of research is focused on financial institutions (see Karminsky and Khromova 2016; Cao et al. 2006; Karminsky and Kostrov 2014) while we found fewer efforts related to credit risk modeling of ICs. The latter explored a great variety of applied models, ranging from simple univariate studies to artificial intelligence (AI) methods. Univariate methods (UM) and discriminant analysis (DA) were very popular until 1990: since that, the other models—such as ordered logistic regression (OLR) and ordered probit regression (OPR), neural network (NN), support vector machine (SVM) and random forest (RF)—have become more widespread due to advancements in technologies (Bellovary et al. 2007). Demeshev and Tikhonova (2014) compared the ability of several linear and non-linear statistical methods (linear, quadratic and mixture DA, OLR, OPR and RF) to predict the default risk of small and medium Russian ICs. He showed that linear algorithms had less prediction power than that of non-linear ones, from which RF demonstrated the highest accuracy. However, Demeshev did not expand his study to large firms and/or ICs from other BRICS jurisdictions.

Among the above-mentioned studies, there is a distinct subset of efforts aimed at modeling of ICRs of industrial companies by re-producing PCRs. Metz and Cantor (2006) developed a UM model that converted ICs' financial metrics to implied ratings, took an appropriate weighted average of them and forecasted the Moody's PCRs. The testing of the model on the PCRs assigned by the US non-financial, non-utility corporates for 1995–1997 demonstrated that its accuracy (around 27%) exceeded that of ordinary linear regression (around 18%) and OPR (around 20%). We note, however, the limited size of the sample and its concentration on ICs from the USA. Karminsky and Polozov (2016) built the PCRs assigned by S&P and Moody's with the OPR. His sample included 215 ICs from 39 countries observed in the period of 2008–2009. The research proved the hypothesis that, in addition to financial ratios, other factors such as the industry, macroeconomic indicators, and the level of maturity of financial markets are significant in PCR modeling.

Depending on the various set of predictors, the model demonstrated the accuracy in the range of 37–43%. The limitations of this study were (1) application of only one statistical method (OPR); and (2) limited timespan for modeling.

In a few researches, the modeling of ICRs was performed with AI methods. Zan et al. (2004) re-constructed PCRs of ICs from Taiwan and the USA. These PCRs were assigned by Taiwan Rating Corporation and S&P respectively. In this effort, the SVM and NN demonstrated the slightly higher accuracy than OPR. The limitation of this study were (1) the application of abridged rating scale (rating classes only); and (2) the usage of the small sample.

Kumar and Bhattacharya (2006) modeled ICRs from the sample of PCRs assigned by Moody's from 2003 to 2004. The sample included 129 ICs from various countries and industries. The authors applied LDA and NN and used only financial variables in the modeling. The study confirmed that NN had a higher accuracy (79%) in comparison to LDA (33%). Yet, the limitation of this paper was (1) limited sample size; (2) the usage of only financial variables in models.

The conclusion is that the most of the efforts described above were focused on developed markets (mainly US) or separate countries (Russia, Korea, Taiwan, etc.). Very few (if any) efforts were focused on prediction of PCRs of ICs from all BRICS countries. The analyzed papers also had limitations such as small sample sizes, the limited set of models and/or explanatory variables used. This paper is aimed on filling these gaps.

3 The data, the variables and the methods

3.1 The data and explanatory variables

For ICRs modeling, we applied the mechanism developed in Grishunin and Suloeva (2016). It uses rating methodologies of Moody's as the framework. Our data set included 221 ICs that at the year-end 2006–2016 had PCRs from Moody's. The set included the following countries: Brazil (71 companies); Russia (61 companies); India (21 companies); China (41 companies) and South Africa (17 companies). The PCRs for these ICs were obtained from Bloomberg. We note however, that for some issuers the PCRs were available for less than 10 years. The companies in the set belonged to 13 distinct industries: oil and gas (20 companies); chemical (9 companies); manufacturing (11 companies); mining (15 companies); utilities and power companies (49 companies); transportation (30 companies); telecommunication (8 companies); steel (13 companies); retail (2 companies); protein and agriculture (8 companies); real estate, building materials and construction (25 companies); paper and forest products (3 companies); business and consumer goods (18 companies). The total number of panel data observation was 1217. The set was divided into a training sample (in-sample; 857 observations) and a validation sample (out-of-sample; 362 observations). We applied Moody's rating scale for ICR system with alpha scores from Ca-Ba3 to Aa3-Aaa (see Table 6 in "Appendix"). This scale consists of 13 grades, each grade is also mapped to a numerical scale from 1 to 13.

The explanatory variables (EVs) included (1) financial variables which reflected the ICs performance; (2) dummy variables for home region, industry, affiliation with the state; and (3) macroeconomic variables in the ICs' country of residence. Financial variables were chosen from Moody's methodologies for non-financial corporations (Moody's 2018). They contained five components: (1) the business profile (2) the size; (3) profitability; (4) the debt leverage and the interest coverage; and (5) the financial policy. Three of them are directly inferred from the companies' financial reporting: the size (revenue), the profitability (the earnings before interest and tax (EBIT) margin); the debt leverage and the interest coverage. The remaining two are evaluated by subjective analysis of companies' business environment. For all components, we selected EVs, which were the best matches taken from Moody's methodologies (see Table 1).

Financial data of IC were obtained from their IFRS or GAAP financial statements and/or annual reports. These statements, in turn, were taken from Capital IQ. We also adjusted financial metrics as required by Moody's methodology (Moody's 2016).

Data for macroeconomic EVs were supplied from World Bank. The list of macroeconomic and dummy EVs is shown in Table 2.

3.2 Statistical methods and the models

3.2.1 Linear discriminant analysis (LDA)

Linear discriminant analysis allows to discriminate between two or more groups of objects by multiple variables at the same time. The goal is to discover the linear combination of variables (the discriminant function LD) that optimally divides the groups in question (Tharwat et al. 2017):

$$LD_{ik} = a_{1k} * x_{i1k} + a_{2k} * x_{i2k} + \cdots + a_{jk} * x_{ijk} + \cdots + a_{mk} * x_{ink},$$

i-the object, k-the group, n-the total number of objects, a_{jk}—coefficients.

Table 1 List of financial explanatory variables

EV's description and notation	Formula and explanation
Revenue (Revenue), \$ million	ICs 12-month gross revenue at the year end
EBIT margin (EBITmargin), %	Ratio of earnings before interest and tax to revenue $Em = \frac{EBIT}{Rev}$
Interest coverage (Eie), x	Ratio that indicates how much times interest is covered by EBIT $Eie = \frac{EBIT}{Interest}$
Gearing ratio (Dbc), x	Calculated as ratio of book value of debt to book value of equity $Dbc = \frac{Debt}{Equity}$
Financial leverage (RCF_d), %	Cash flow debt coverage $RCF\_d_i = \frac{OCF-CWC-Dividend}{Debt}$ OCF—operating cash flow of IC CWC—change in working capital

Table 2 List of macroeconomic and dummy variables

EVs description and notations	Formula and explanations
Dummy variables	
IC is located in Russia (RK)	1—if IC is Russia-based; 0—if opposite
IC is located in China (China)	1—if IC is China-based, 0—if opposite
IC is supported by the government (Rtg)	1—if IC is a government owned or significantly controlled entity, 0—if opposite
IC is operating in certain industry: Oil and gas (Og) Chemical (Ch) Utilities and power generation (UaPC) Transportation (Tran) Telecommunication (Tele) Retail (Retail) Protein and agriculture (PA) Real estate and construction (Re) Paper and forest products (PFP) Manufacturing (Man) Business and consumer goods (BaCGaS) Steel (Steel)	1—if operates in given industry, 0—if opposite
Macroeconomic variables	
GDP per capita (GDPpc), $	Gross domestic product (GDP) per capita in current $
Inflation (Infl), %	Consumer price index (% to previous year)
Exports to GDP (Exp), %	The ratio of export to gross domestic product (% of GDP)

LDA assumes that the descriptions of objects of each K-th class are the manifestation of the multidimensional random variable distributed normally $N_m(\mu_k; \Sigma_k)$. Therefore, p linear discriminant functions must be found, p will be equal minimum of (1) the number of sets minus 1; or (2) the number of EVs. The criterion for calculation of coefficient of discriminant function is: the better the classification of EVs, the smaller the scattering of points relative the centroid within the group and the greater the distance between the centroid of the groups.

After LDs have been constructed, it is possible to classify any observation by inserting values of EVs in discriminant equations for each k-th group and calculate the response values, $k = 1, 2, \dots, p$. The results of ICR modeling with LDA are in Table 3 and Fig. 1.

Table 3 Proportions of trace of each discriminant functions

Discriminant function	Proportion of trace, %
LD1	0.6995
LD2	0.0776
LD3	0.0579
……	…
LD12	0.0013

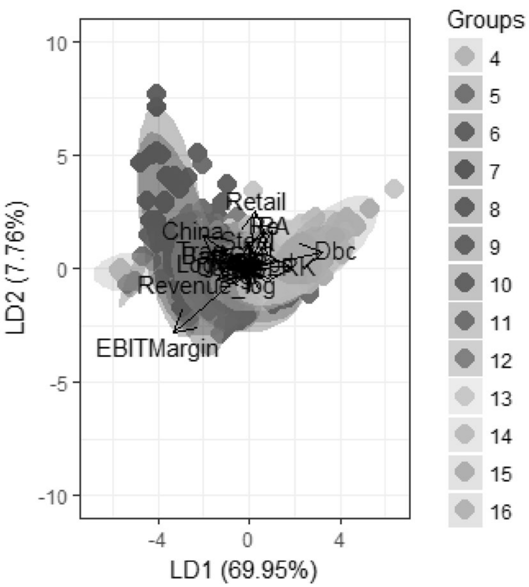

Fig. 1 Visualization of results of classification with LD1 and LD2

3.2.2 Ordered logistic regression (OLR)

In this model, for k-ordered alternatives (ICRs mapped in numerical scale), the probability that IC with the number m and the set of EVs Y_m will be classified in grade k, equals (Karminsky and Peresetsky 2007):

$$\begin{cases} P(Y_m = 1) = F(c_1 - Y_m\beta) \\ P(Y_m = 2) = F(c_2 - Y_m\beta) - F(c_1 - Y_m\beta) \\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots \\ P(Y_m = k-1) = F(c_{k-1} - Y_m\beta) - F(c_{k-2} - Y_m\beta) \\ P(Y_m = k) = 1 - F(c_{k-1} - Y_m\beta) \end{cases}.$$

The function F is the logistic distribution function. The model's parameters are the vector of coefficients (β) and vector of thresholds $c=(c_1, c_2,\ldots,c_{k-1})$. These parameters are estimated with method of maximum likelihood with Huber-White standard errors. Application of the OLR gave the following results (Table 4).

Based on the results of the likelihood ratio test, LR = 1359.90. This value is above $\chi^2(q)$ by 99%; therefore, the hypothesis that the model is statistically insignificant is rejected. Another statistical measure of quality, Pseudo R^2 is ensured by the level of 66.9%. This Pseudo R^2 is higher than that achieved in Karminsky (2011)—0.399 or lower.

Table 4 Result of ICR modeling with OLR

Variable notation	Coefficient	Standard error	t-criteria
BaCGaS	−0.63	0.355	−1.77*
Ch	−0.52	0.396	−1.32
Man	0.18	0.441	0.42
Og	−0.43	0.324	−1.32
PFP	0.44	0.463	0.95
PA	2.05	0.439	4.67***
Re	1.56	0.365	4.27***
Retail	1.53	0.694	2.20**
Steel	0.62	0.336	1.84*
Tele	−0.65	0.399	−1.63*
Tran	−2.05	0.345	−5.93***
UaPC	−0.54	0.295	−1.83*
Rtg	−0.90	0.176	−5.08***
China	−2.74	0.315	−8.68***
RK	3.00	0.317	9.44****
Lg(GDPpc)	−1.12	0.522	−2.14**
Infl	0.01	0.030	0.32
Exp	−0.05	0.014	−3.72***
Lg(Revenue)	−2.15	0.144	−14.91***
EBITMargin	−5.11	0.408	−12.52***
Eie	−0.03	0.008	−4.07***
Dbc	5.07	0.459	11.04***
RCF_d	0.81	0.235	3.45***
LR Chi^2	1359.90		
Degrees of freedom	23		
$P(L > Chi^2)$	< 0.0001		
Pseudo R^2	0.68		

***, **, *—The coefficient is significant at levels of 1, 5, and 10%, respectively

3.2.3 Support vector machine (SVM)

Support vector machine uses a linear model to implement non-linear class boundaries through the mapping and input vectors into a high-dimensional feature space. The linear model constructed in the new space can represent a nonlinear decision boundary in the original space. In the new space, an optimal separating hyperplane (OSH) is constructed. Thus, with SVM, a special linear model—the maximum margin hyperplane (MMH), which gives the maximum segregation between decision classes—can be found. The training examples that are in the critical zone, closest to MMH are called support vectors. All other training examples are irrelevant for defining the class boundaries (Lee 2007).

However, the standard SVM formulation solves only the binary classification problem and cannot be transferred for the cases that require classification of object to multiple grades (as required for ICR modeling). To account for non-linearity and multiple grades, the variable space is extended with the special kernel function, allowing the building of models that use separating hyperplanes of various forms (Hájek and Olej 2011).

To construct SVM, we applied "one-against-one" as it proved to be an effective method for solving problems of rating forecasting (Zan et al. 2004) We also used the kernel with radial basis function (RBF). Then, the OSH will be computed by the selection of coefficient α_i in:

$$z_k(x) = \sum_{i=1}^{p} \alpha_i exp\left(\gamma \left\| x_i - x_j \right\|^2\right) + \beta_0,$$

where, p—dimension; x_i,x_j—vectors; γ, β_0—parameters of RBF.

To solve for α_i, quadratic optimization using Lagrange multipliers is used. We also used $\gamma = 0.5$.

3.2.4 Artificial neural network (NN)

We applied three-layer fully connected backpropagation NN (Zan et al. 2004). The input layer nodes are EVs, output nodes are modeled ICR and the number of hidden layer nodes is (the number of input nodes + number of output nodes)/2. Activations flow started from the input layer via the hidden layer and then to the output layer. The architecture of NN is presented at the Fig. 2.

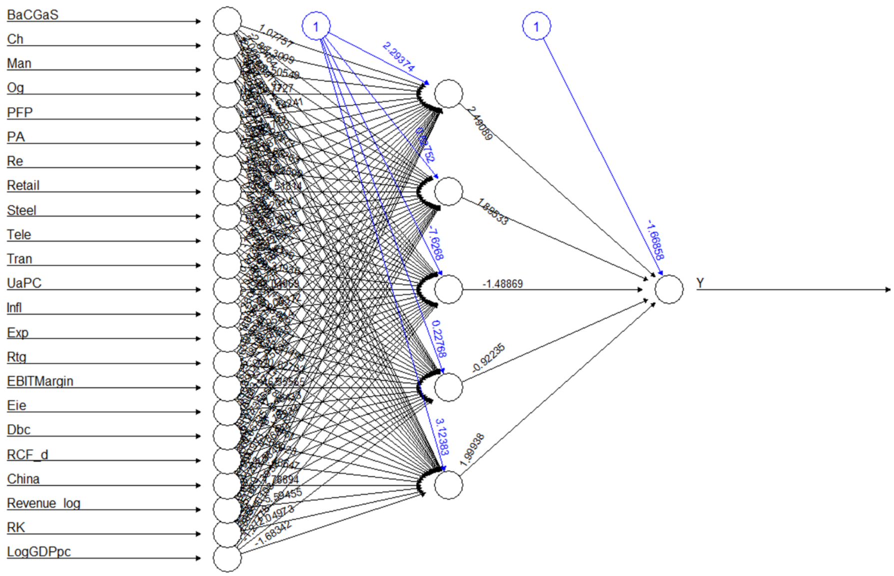

Fig. 2 The visualization of backpropagation NN used for ICR modeling

We trained our NN with the function neural net in R. In this function, the training starts with a random set of weights, the weights are adjusted each time NN sees the input–output pairs that are processed via the forward pass and the backward pass. During the latter, the NN's achieved output is compared with the target output and errors are computed comparatively to the input units. To reduce the errors, the weights connected to the output units are adjusted (a gradient descent). The network adjusts its weights incrementally until the NN stabilizes.

3.2.5 Random forest (RF)

Random forest consists of a collection or ensemble of simple tree predictors, each capable of producing a response when presented with a set of predictor values. RF is a deep statistical method, which performs bootstrap aggregation for a set of trees of decisions. When constructing each individual tree (we built 500 trees), some of the observations will not be used, and some of the observations will be used several times. In the algorithm, there is random selection from observations with iterations from the original sample set. To construct each tree split, the random selection of the number of regressors from the whole set of regressors is performed (we used 3 regressors) and then the best criteria from them which gives the largest decrease in Gini criterion is selected. This construction approach corresponds to the key principle of ensemble learning—the algorithms must be accurate and diverse (so each tree is built on its own training sample and in selection of each split there is an element of chance). Studies show that its advantages include high prediction accuracy, avoidance of over-fit and robustness against high dimensional data (Saitoh 2016).

In the modeling of ICRs the predicted RF result is determined based on the average output value of the plurality of regression trees. The value predicted by the RF is calculated:

$$\hat{y}(x_i) = \frac{1}{B}\sum_{b=1}^{B} h(x_i;\, T_b),$$

where x_i is the i-th attribute data, B is the number of regression trees, h is the output of regression tree T_b.

The accuracy criteria of value predicted is the estimation of probability of classification error of random forest in the confusion matrix of the prediction. This estimation is done by out-of-bag of performance (OOB) method. The training sample consists of 2/3 of input objects, the remaining set consists of 1/3 of input objects (OOB). The sum of square errors (SSE) is calculated at each split point between the predicted value $\hat{y}(x_i)$ and the actual values. The variable resulting in minimum SSE is selected for the node. Then, this process is recursively continued till the entire data is covered. The mean SSE is used for evaluation of accuracy of prediction in the confusion matrix:

$$MSE \sim MSE^{OOB} = n^{-1}\sum_{i=1}^{n}(\hat{Y}^{OOB}(X_i) - Y_i)^2.$$

The most significant EVs in RF model are presented in Fig. 3.

3.3 Measuring the accuracy of ICR models

The validation of the accuracy of the model reflects the hit rate (HR) of modeled ICR to actual PCR, i.e.

$$HR = \frac{\sum_{i=1}^{M} x_i}{M},$$

x_i—the binary variable, equals 1 if the modeled ICR hits actual PCR, 0 if opposite, M—the number of observation in the validation sample.

Consequently, the modeling errors can be evaluated by the accuracy ratios (AR):

$$AR_{\Delta=Z} = \frac{\sum_{j=1}^{N} w_{\delta=\Delta \cdot j}}{M},$$

$w_{\delta=\Delta,j}$—the binary variable, equals 1 if the modeled error $\Delta=$PCR-ICR equals Z. 0 if opposite. In the same way, we can calculate $AR|_{\Delta}|_{\leq Z.}$

The accuracy of the model can be also characterized by the Type I and Type II errors. Type I errors are overstatement of modeled ICRs in comparison to actual PCRs. Type II errors are reverse—understatement of modeled ICR in comparison to actual PCRs. It is generally agreed upon that Type I errors are costlier than Type II errors for several reasons including loss of business, damage to a firm's reputation and potential lawsuits (Bellovary et al. 2007). Therefore, the model, which results in fewer Type I errors compared to Type II errors, is considered the best among the alternative models.

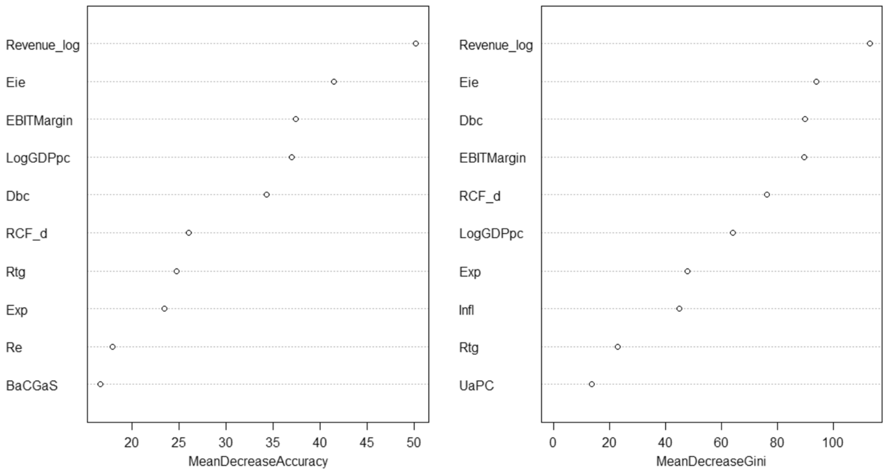

Fig. 3 The visualization of the most significant EVs in RF method

4 Results

The outcome of ICR modeling with above-mentioned statistical methods and its comparison to actual PCRs is presented in the Table 5. Negative Δ represents Type I error while positive Δ gives Type II errors.

The findings are:

1. AI methods (SVM, NN and RF) outperform "traditional" modeling methods (LDA and OLR) by predictive power. On the training sample, AI gives hit rate of 67.5% on average and 54.5% on average under the out-of-sample fit check. Conversely, LDA and OLR, if considered together, give hit rate of only 41.7% on average on training sample and only 37.4% on validation sample.
2. AI methods also outperform LDA and OLR by smaller error spread. In comparison to LDA and OLR, which give maximum error of 2–3 notches from actual PCRs, AI methods demonstrate very small percentage of errors above 1 notch (RF gives none).
3. AI performs better than traditional methods by the distribution of Type I and Type II errors. Unlike that in OLR and LDA, which give nearly symmetrical Type I and Type II errors, the number of Type I errors in AI model outcomes is very small and do not exceed 2.5% in total.
4. The results show the slight deteriorations in the predictive power of the models under the out-of-sample fit check. This level deterioration is expected (Karminsky 2011). However, the accuracy of RF-based model deteriorates materially (to 58% on validation sample from 100% on training sample). Additional research is necessary for turning the algorithm to limit such deterioration.
5. Among the "traditional" methods, under the out-of-sample fit check, LDA slightly outperforms OLR by the predictive power (39.7% vs. 35%). These hit rates are comparable with those reported in Karminsky (2011) of 38.8–41.9%.

Table 5 The outcome of ICR modeling and its comparison to PCRs of BRICS' industrial companies

Model	Sample	Hit rate and accuracy ratios, %						
		$\Delta = -2$	$\Delta = -1$	$\Delta = 0$	$\Delta = 1$	$\Delta = 2$	$\lvert\Delta\rvert \le 1$	$\lvert\Delta\rvert \le 2$
LDA	In-sample	7.1	15.0	45.2	13.9	8.9	74.8	90.0
	Out of sample	11.6	16.9	39.7	12.3	6.3	68.8	86.8
OLR	In-sample	8.8	18.6	38.1	18.7	6.7	75.6	91.0
	Out of sample	5.6	20.7	35.0	15.9	9.8	71.7	87.0
SVM	In-sample	0	1.8	47.6	49.1	1.5	98.5	100
	Out of sample	0	0.3	54.2	44.4	1.1	98.9	100
NN	In-sample	0	1.9	55.0	41.4	1.7	98.3	100
	Out of sample	0	2.4	51.4	44.3	1.9	98.0	100
RF	In-sample	0	0	100	0	0	100	100
	Out of sample	0	2.3	58	39.7	0	100	100

6. Consequently, among the AI methods, under the out-of-sample fit check, RF gives the highest accuracy (58%) followed by SVM (54%) and NN (51%). Additionally, for RF, in 100% of the cases, the prediction error does not exceed 1 notch (for SVM and NN—in almost 99% cases).
7. We must mention, however, that AI models are "black boxes" because they cannot provide easy interpretation, of which EVs are the most significant. This feature may limit the practical application of these models.

5 Conclusion

This paper is devoted to the comparison of the ability of various statistical methods to reproduce PCRs of BRICS' industrial companies using publicly available information. This topic is important because a lot of these companies lack PCRs from reputable CRAs and investors must model the ICRs as the proxies of PCRs. We compared the performance of the five statistical methods (linear discriminant analysis (LDA), ordered logit regression (OLR), support vector machine (SVM), neural network (NN) and random forecast (RF)) in reconstruction of Moody's PCRs of 208 industrial companies in 2006–2016. The resulting models were checked for in-sample and out-of-sample predictive fit.

Among considered methods, AI models (SVM, NN and RF) outperformed LDA and OLR by (1) predictive fit; and (2) distribution of Type I and Type II errors. On the validation sample, AI methods gave hit rate of 54.5% on average and 99% of modeled ICRs predicted actual PCRs with the errors not exceeding 1 notch. Consequently, LDA and OLR gave hit rate of only 37.4 on average and only 70.2% of modeled ICRs predicted actual PCRs with the errors not exceeding 1 notch. Unlike in OLR and LDA, which give symmetrical Type I and Type II error, the share of Type I errors in models produced by AI is very small and do not exceed 2.5%. We can, therefore, conclude that AI methods should have a significant practical use for predicting the PCRs of ICs from BRICS countries.

Appendix: Internal credit rating scale, descriptive statistics and inter-factor correlation

See Tables 6, 7 and 8.

Table 6 Rating scale of modeled internal credit ratings

Rating grade	Aa3-Aaa	A1	A2	A3	Baa1	Baa2	Baa3
Numerical scale	1	2	3	4	5	6	7
1-year default rate[a], %	0	0.07	0.05	0.05	0.13	0.17	0.25

Rating grade	Ba1	Ba2	Ba3	B1	B2	Ca-B3
Numerical scale	8	9	10	11	12	13
1-year default rate, %	0.44	0.71	1.36	1.97	2.95	12.9

Source: Moody's (2017)

[a]Average 1-year default rate calculated in 1983–2017 by Moody's (2017)

Table 7 Descriptive statistics of model's financial variables

Variables	Notation	Average	Maximum	Minimum	Standard deviation
Inflation, %	Infl	6.75	14.12	1.44	2.87
Share of export of GDP, %	Exp	21.5	46.5	10.7	8.20
EBIT margin, %	EBITmargin	24.6	164.4	−75	20.3
Operating profit/interest expenses, x	Eie	8.2	228.1	−4.2	19.6
Debt/Book capitalization, x	Dbc	0.48	1.47	0.02	0.20
Retained cash flow (RCF)/Debt, %	RCF_d	34.5	112.7	−58.9	71.9
Lg(Revenue), x	Revenue_log	6.55	8.66	4.19	0.74
Lg(GDP per capita), x	LogGDPpc	4.17	4.41	3.54	0.19

Table 8 Matrix of inter-factor correlation of model's financial variables

	Infl	Exp	EBIT margin	Eie	Dbc	RCF_d	Revenue_log	Log GDPpc
Infl	1.00							
Exp	0.25	1.00						
EBIT margin	−0.03	−0.15	1.00					
Eie	0.04	0.15	0.09	1.00				
Dbc	−0.09	−0.35	0.07	−0.43	1.00			
RCF_d	0.10	0.19	0.02	0.84	−0.49	1.00		
Revenue_log	−0.04	0.23	−0.44	0.16	−0.25	0.13	1.00	
LogGDPpc	0.16	0.09	0.03	−0.20	−0.02	−0.11	−0.12	1.00

References

Bellovary, J. L., Giacomino, D. E., & Akers, M. D. (2007). A review of bankruptcy prediction studies: 1930 to present. *Journal of Financial Education, 33,* 1–42.

Cao, L., Guan, L. K., & Jingqing, Z. (2006). Bond rating using support vector machine. *Intelligent Data Analysis, 10,* 285–296.

Demeshev, B., & Tikhonova, A. (2014). Default prediction for Russian companies: Intersectoral comparison. *Economic Journal of Higher School of Economics, 18*(3), 359–386.

Grishunin, S., & Suloeva, S. (2016). Development of the credit risk assessment mechanism of investment projects in telecommunication. In O. Galinina, S. Balandin, & Y. Koucheryavy (Eds.), *NEW2AN/ruSMART 2017* (Vol. 9870, pp. 752–765). Russia: LNCS.

Hájek, P., & Olej, V. (2011). Credit rating modelling by kernel-based approaches with supervised and semi-supervised learning. *Neural Computing and Applications, 20*(6), 761–773.

Jarrow, R. (2009). Credit risk models. *Annual Review of Financial Economics, 1,* 37–68.

Karminsky, A. (2011). Corporate rating models for emerging markets. *Journal of Corporate Finance Research, 5*(3), 19–29.

Karminsky, A., & Khromova, E. (2016). Modelling banks' credit ratings of international agencies. *Eurasian Economic Review, 6,* 341–363.

Karminsky, A., & Kostrov, A. (2014). The probability of default in Russian banking. *Eurasian Economic Review, 4*(1), 81–98.

Karminsky, A., & Peresetsky, A. (2007). Models of ratings of international rating agencies. *Prikladnaia Ekonometrika (Applied Econometrics), 1,* 3–19. **(in Russian)**.

Karminsky, A., & Polozov, A. (2016). *Handbook of ratings: Approaches to ratings in the economy, sports, and society* (Vol. 338). Berlin: Springer.

Kumar, K., & Bhattacharya, S. (2006). Artificial neural network vs linear discriminant analysis in credit ratings forecast: A comparative study of prediction performances. *Review of Accounting and Finance, 5,* 216–227.

Lee, Y. C. (2007). Application of support vector machines to corporate credit rating prediction. *Expert Systems with Applications, 33,* 67–74.

Metz, A., & Cantor, R. (2006). Moody's Credit Rating Prediction Model. Special Comment. Moody's Investor Service. https://www.moodys.com/sites/products/DefaultResearch/2006200000425644.pdf.

Moody's. (2016). Financial statement adjustments in the analysis of non-financial corporations. Moody's Investor Service. https://www.moodys.com/research/Financial-Statement-Adjustments-in-the-Analysis-of-Non-Financial-Corporations–PBC_1050255.

Moody's. (2017). Cross-sector: The performance of Moody's corporate debt ratings—2017Q2. Moody's Investor Service. https://www.moodys.com/researchdocumentcontentpage.aspx?docid=PBC_1084629.

Moody's. (2018). List of rating methodologies. Moody's Investor Service. https://www.moodys.com/research/List-of-Rating-Methodologies–PBC_127479.

Saitoh, F. (2016). Predictive modeling of corporate credit ratings using a semi-supervised random forest regression. In *IEEE international conference on industrial engineering and engineering management* (pp. 429–433).

Staples, D., et al. (2013). Credit considerations for corporates in emerging markets. Moody's Investor Services. http://www.moodys.com/research/Credit-Considerations-for-Corporates-in-Emerging-Markets–PBC_153812.

Tharwat, A., Gaber, T., Ibrahim, A., & Hassanien, A. E. (2017). Linear discriminant analysis: A detailed tutorial. *AI Communications, 30,* 169–190.

Zan, H., Chen, H., Hsu, C.-J., Chen, W.-H., & Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: A market comparative study. *Decision Support Systems, 37,* 543–558.

Affiliations

Alexander M. Karminsky[1] · Sergei Grishunin[2] · Natalya Dyachkova[1] · Maxim Bisenov[3]

Alexander M. Karminsky
akarminsky@hse.ru

Sergei Grishunin
sergei.v.grishunin@gmail.com

Maxim Bisenov
Max.Bisenov@yandex.ru

[1] Faculty of Economic Science, School of Finance, Higher School of Economics, 26-4 Shabolovka Street, Moscow, Russia

[2] Peter the Great St. Petersburg Polytechnic University, 27 Polytechnicheskaya Street, St. Petersburg, Russia

[3] Faculty of Mathematical Methods in Economy, Plekhanov Russian University of Economics, 36 Stremyanny Lane, Moscow, Russia