Eurasian Economic Review

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

--Manuscript Draft--

Manuscript Number:	EAER-D-18-00092R1
Full Title:	The comparison of empirical methods for modeling credit ratings of industrial companies from BRICS countrie
Article Type:	Research Paper
Funding Information:	
Abstract:	We compared the ability of various empirical methods to reproduce public credit ratings (PCRs) of industrial companies (ICs) from BRICS countries using pub-licly available information. This task is of primary importance for researchers and practitioners as a lot of BRICS' ICs lack public credit ratings (CRs) from reputa-ble rating agencies such as Moody's, Standard and Poor's or Fitch. The paper is aimed at filling the gap in the existing research as only very few efforts were fo-cused on prediction of PCRs of ICs from entire BRICS IC community. The modelled variables are CRs of 208 BRICS' industrial companies assigned by Moody's at the year-end from 2006 till 2016. The sample included 1217 observa-tions. Financial dependent variables included companies' revenue, operating prof-itability, interest coverage ratio, debt/book capitalization and cash flow debt cov-erage. Non-financial dependent variables included dummies for home region, in-dustry, 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 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 modelling, Industrial company, Linear discriminant analysis, Ordered logit model, Artificial intelligence methods JEL classification: C50-C53, G31-G33
Corresponding Author:	Natalya Fedorovna Dyatchkova, M.D. Nacional'nyj issledovatel'skij universitet Vyssaa skola ekonomiki RUSSIAN FEDERATION
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Nacional'nyj issledovatel'skij universitet Vyssaa skola ekonomiki
Corresponding Author's Secondary Institution:	
First Author:	Alexander Karminsky, D.S.
First Author Secondary Information:	
Order of Authors:	Alexander Karminsky, D.S.
	Sergei Grishunin
	Natalya Fedorovna Dyatchkova, M.D.
	Maxim Bisenov
Order of Authors Secondary Information:	

Author Comments:	I am writing to submit our manuscript entitled, "The comparison of empirical methods for modeling of credit ratings of industrial companies from BRICS countries" for consideration as an article for Eurasian Business Review journal. Our research was presented at EBES conference in Berlin. Should you select our manuscript for peer review, we would like to suggest the following potential reviewers/referees because they would have the requisite background to evaluate our findings and interpretation objectively. Each named author has substantially contributed to conducting the underlying research and drafting this manuscript. Sincerely, Natalya
Response to Reviewers:	The literature review section was expanded, the most recent papers were also added and analyzed. The review was also linked with theoretical background. The motivation for selecting BRICS countries was added in the introduction and literature review. The name of the variable was changed as advised. The list of dependent variables was clarified in section 3.2. The name of the variable was changed as advised. Section 3.3 was re-writen to remove the confusion with deltas. The results were segregated from the models explanations. Necessary explanation was added explaining why certain deltas are used. Findings are presented as normal text. Explanation why BRICS countries were added to the text. We added several paragraphs in secttion 3 to explain the motivation of selecting BRICS.

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

Alexander M. Karminsky[1], Sergei Grishunin[2], Natalya Dyatchkova[3] and Maxim Bisenov[4]

[1] Professor, Faculty of Economic Science/ School of Finance, Higher School of Economics,
26-4 Shabolovka Street, Moscow, Russia
akarminsky@hse.ru
[2] Candidate of Econ. Sciences, Peter the Great St. Petersburg Polytechnic University,
27 Polytechnicheskaya Street, St. Petersburg, Russia
sergei.v.grishunin@gmail.com
[3] Faculty of Economic Science/ School of Finance, Higher School of Economics,
26-4 Shabolovka Street, Moscow, Russia
Nfdyachkova@gmail.com
[4]Faculty of mathematical methods in economy, Plekhanov Russian University of Economics,
36 Stremyanny lane, Moscow, Russia
Max.Bisenov@yandex.ru

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 of primary importance for researchers and practitioners as a lot of BRICS' ICs lack public credit ratings (CRs) from reputable rating agencies such as Moody's, Standard and Poor's or Fitch. The paper is aimed at filling the gap in the existing research as only very few efforts were focused on prediction of PCRs of ICs from entire BRICS IC community. The modelled variables are CRs of 208 BRICS' industrial companies assigned by Moody's at the year-end from 2006 till 2016. The sample included 1217 observations. Financial dependent variables included companies' revenue, operating profitability, interest coverage ratio, debt/book capitalization and cash flow debt coverage. Non-financial dependent 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 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 modelling, Industrial company, Linear discriminant analysis, Ordered logit model, Artificial intelligence methods

JEL classification: C50-C53, G31-G33

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

Alexander M. Karminsky[1], Sergei Grishunin[2], Natalya Dyatchkova[3] and Maxim Bisenov[4]

[1] Professor, Faculty of Economic Science/ School of Finance, Higher School of Economics, 26-4 Shabolovka Street, Moscow, Russia
akarminsky@hse.ru
[2] Candidate of Econ. Sciences, Peter the Great St. Petersburg Polytechnic University, 27 Polytechnicheskaya Street, St. Petersburg, Russia
sergei.v.grishunin@gmail.com
[3] Faculty of Economic Science/ School of Finance, Higher School of Economics, 26-4 Shabolovka Street, Moscow, Russia
Nfdyachkova@gmail.com
[4]Faculty of mathematical methods in economy, Plekhanov Russian University of Economics, 36 Stremyanny lane, Moscow, Russia
Max.Bisenov@yandex.ru

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 of primary importance for researchers and practitioners as a lot of BRICS' ICs lack public credit ratings (CRs) from reputable rating agencies such as Moody's, Standard and Poor's or Fitch. The paper is aimed at filling the gap in the existing research as only very few efforts were focused on prediction of PCRs of ICs from entire BRICS IC community. The modelled variables are CRs of 208 BRICS' industrial companies assigned by Moody's at the year-end from 2006 till 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 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 modelling, Industrial company, Linear discriminant analysis, Ordered logit model, Artificial intelligence methods

JEL classification: C50-C53, G31-G33

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

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 of primary importance for researchers and practitioners as a lot of BRICS' ICs lack public credit ratings (CRs) from reputable rating agencies such as Moody's, Standard and Poor's or Fitch. The paper is aimed at filling the gap in the existing research as only very few efforts were focused on prediction of PCRs of ICs from entire BRICS IC community. The modelled variables are CRs of 208 BRICS' industrial companies assigned by Moody's at the year-end from 2006 till 2016. The sample included 1217 observations. Financial dependent variables included companies' revenue, operating profitability, interest coverage ratio, debt/book capitalization and cash flow debt coverage. Non-financial dependent 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 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 modelling, Industrial company, Linear discriminant analysis, Ordered logit model, Artificial intelligence methods

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 forecas the creditworthiness of these assets.

One of the tools which 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 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 2016).

Still, 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 modelling also helps to quick replication of the model.

The well-proven method of modelling 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: 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 modelling; 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); (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 modelling. Section 3 explores the data, the set of explanatory variables and the methods of modelling. Finally, in section 4, the accuracy of each method of reproducing Moody's PCRs is discussed and conclusions are formulated.

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

Majority of efforts points out the growing magnitude of credit risks in business and importance of its modelling 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 less number of efforts related to credit risk modelling of ICs. The latter explored a great variety of applied models starting from simple univariate studies to artificial intelligence 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 and Giacomino 2007). Demeshev (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 default 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 above-mentioned studies, there is a distinct subset of efforts aimed at modelling of ICRs of industrial companies by re-producing PCRs. Metz and Cantor (Metz and Cantor 2006) worked out the 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. In turn, Karminsky (Karminsky 2015) re-constructed the PCRs assigned by S&P and Moody's with the OPR. His sample included 215 ICs from 39 countries observed in 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 modelling. Depending on the various set of predictors, the model demonstrated the accuracy in range of 37%-43%. The limitations of this study were (1) application of only one statistical method (OPR); and (2) limited timespan for modelling.

In a few researches the modelling of ICRs was performed with artificial intelligent methods. Zan et al. re-constructed PCRs of ICs from Taiwan and the USA. These PCRs were assigned by Taiwan Rating Corporation and S&P respectively. (Zan et al. 2004). In this effort, the SVM and NN demonstrated the slightly higher accuracy than OPR. The limitationв 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 modelled ICRs from the sample of PCRs assigned by Moody's in 2003-2004. The sample included 129 ICs from various countries and industries (Kumar and Bhattacharya 2006). The authors applied LDA and NN and used only financial variables in the modelling. 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 dependent and explanation variables

For ICRs modelling we applied the mechanism developed in (Grishunin and Suloeva 2016). Its uses rating methodologies of Moody's as the framework. Our data set included 221 IC which 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 in Appendix). This scale consists of 13 grades, each grade is also mapped to a numerical scale from 1 to 13 and to idealized default probabilities.

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 match Moody's methodologies (Table 1).

Table 1. **List of financial explanatory variables**

EV's description and notation	Formula and explanation
Revenue (Revenue), $ million	IC"s 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 which 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 $RCF\_d = \frac{OCF - CWC - Dividend}{Debt}$ OCF – operating cash flow of IC CWC – change in working capital

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.

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)	1 – if operates in given industry, 0 – if opposite

• Protein and agriculture (PA) • Real estate and construction (Re) • Paper and forest products (PFP) • Manufacturing (Man) • Business and consumer goods (BaC-GaS) • Steel (Steel)	
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)

3.2 Statistical methods and the modelling process

Linear discriminant analysis (LDA)

LDA allows to discriminate between two or more groups of objects by multiple variables at the same time. The goal is the discovering such 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

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, $\boldsymbol{k = 1, 2, \dots, p}$. The results of ICR modelling with LDA are in Table 3 and Figure 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

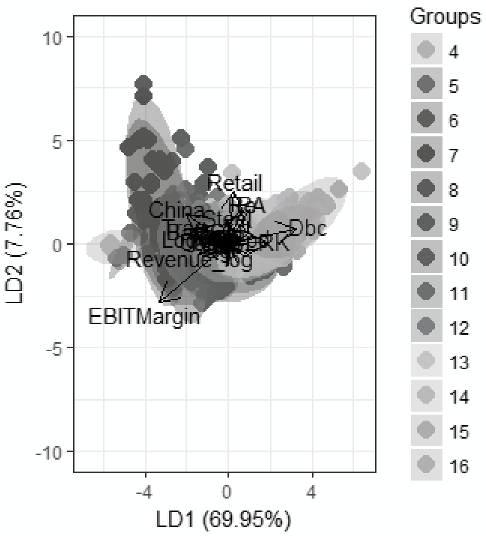

Fig. 1. Visualization of results of classification with LD1 and LD2

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 Polozov 2016):

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

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

Table 4. Result of ICR modelling 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**		

Note: ***, **, * - the coefficient is significant at levels of 1, 5, 10% respectively

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 2015) – 0.399 or lower.

Support Vector Machine (SVM)

SVM uses a linear model to implement non-liner class boundaries through nonlinear mapping 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 can be found the maximum margin hyperplane (MMH) which gives the maximum segregation between decision classes. The training example which are lying 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 which require classification of object to multiple grades (as required for ICR modelling). To account for non-linearity and multiple grades, the variable space is extended with the special kernel function. This allows to build the models with usage of separating hyperplane of various form (Ha´jek 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 γ=0.5.

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 modelled 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.

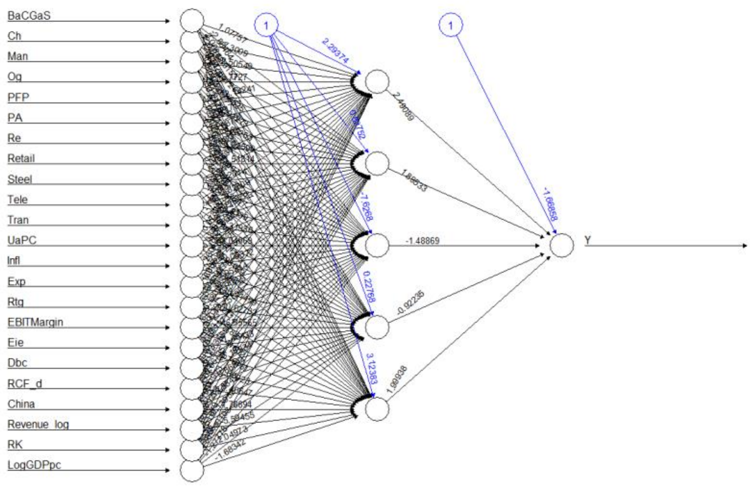

Fig. 2. The visualization of backpropagation NN used for ICR modelling

We trained our NN with the function neuralnet 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 which 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 the output units. To reduce the errors, the weights connected to the output units are adjusted to reduce the errors (a gradient descent). The network adjusts its weights incrementally until the NN stabilizes.

Random Forest (RF)

RF 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 performs bootstrap aggregation of a set of a decision trees. 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 repetitions 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). The studies showed that its advantages high prediction accuracy, avoidance of over-fit and robustness against high dimensional data (Saitoh 2016).

In the modelling 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 and 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 o 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.

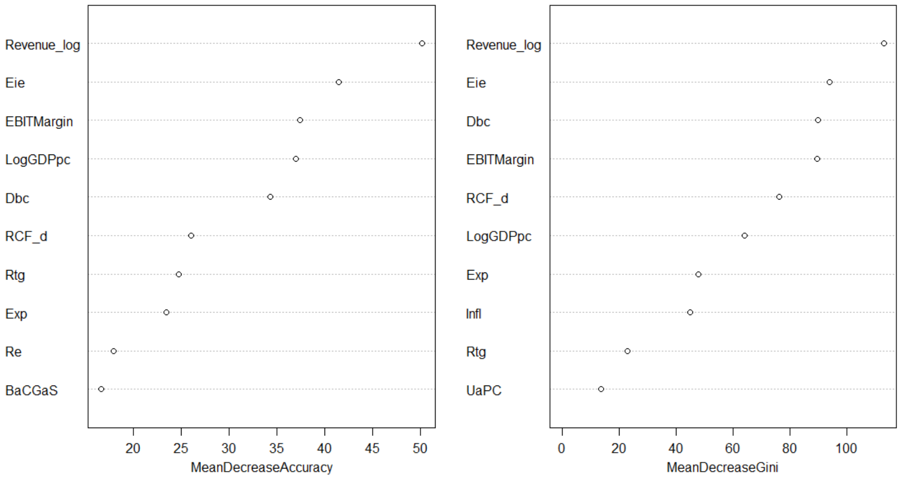

Fig. 3. The visualization of the most significant EVs in RF method

3.3 Measuring the accuracy of ICR model

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

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

x_i – the binary variable, equals 1 if the modelled ICR hits actual PCR, 0 if opposite

M- the number of observation in the validation sample

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

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

$w_{\delta=\Delta,j}$ – the binary variable, equals 1 if the modelled error Δ=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 modelled ICRs in comparison to actual PCRs. Type II errors are reverse – understatement of modelled 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 and Giacomino 2007). Therefore, the model which result in less Type I errors in relation to Type II error among the alternative models to be considered the best.

4 Results

The outcome of ICR modelling 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 error.

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

Model	Sample	Hit rate and accuracy ratios, %						
		Δ=-2	Δ=-1	Δ=0	Δ=1	Δ=2	│Δ│≤1	│Δ│≤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

The findings are:

1. Artificial intelligence (AI) methods (SVM, NN and RF) outperform "traditional" modelling 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 error, 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 2015). 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 2015) of 38.8%-41.9%.
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% 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 which EVs are the most significant. This feature may limit the practical application of these models.

5 Conclusion

This paper is devoted to comparison of the ability of various statistical methods to reproduce PCRs of BRICS's 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 artificial intelligence 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 modelled 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 modelled ICRs predicted actual PCRs with the errors not exceeding 1 notch. Unlike that 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 industrial companies from BRICS countries.

References

1. 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.
2. Cao, L., Guan, L. K., & Jingqing, Z. (2006). Bond rating using support vector machine. Intelligent Data Analysis, 10, 285–296.
3. Cross-sector: the performance of Moody's corporate debt ratings – 2017Q2 (2017). Moody's Investor Service. URL: https://www.moodys.com/researchdocumentcontentpage.aspx?docid=PBC_1084629
4. Demeshev, B., Tikhonova A. (2014). Default Prediction for Russian Companies: Intersectoral Comparison. Economic Journal of Higher School of Economics, 18(3), 359-386
5. Financial statement adjustments in the analysis of non-financial corporations (2016). Moody's Investor Service. URL: https://www.moodys.com/research/Financial-Statement-Adjustments-in-the-Analysis-of-Non-Financial-Corporations--PBC_1050255
6. Grishunin, S., Suloeva, S.: Development of the credit risk assessment mechanism of investment projects in telecommunication (2016). In: Galinina O., Balandin S., Koucheryavy Y. (eds) NEW2AN/ruSMART 2017. LNCS, vol 9870, 752-765.
7. Ha´jek, 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.

8. Jarrow R. (2009). Credit risk models. Annual Review of Financial Economics, 1, 37-68.
9. Karminsky, A. (2015). Corporate ratings models for emerging markets. Corporate Finance, 3, 19-29.
10. Karminsky, A., Khromova E. (2016). Modelling banks' credit ratings of international agencies. Eurasian Economic Review, 6, 341-363.
11. Karminsky, A., Kostrov, A. (2014). The probability of default in Russian banking. Eurasian Economic Review, 4(1), 81–98.
12. Karminsky, A., Peresetsky, A. (2007). Models of ratings of international rating agencies. Prikladnaia Ekonometrika (Applied Econometrics), 1, 3–19. (in Russian).
13. Karminsky, A., Polozov, A. (2016): Handbook of Ratings: Approaches to Ratings in the Economy, Sports, and Society, Springer, 338.
14. 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.
15. Lee, Y. C. (2007). Application of support vector machines to corporate credit rating prediction. Expert Systems with Applications, 33, 67–74.
16. List of rating methodologies (2018). Moody's Investor Service. URL: https://www.moodys.com/research/List-of-Rating-Methodologies--PBC_127479
17.Metz, A., Cantor, R. (2006). Moody's Credit Rating Prediction Model. Special Comment. Moody's Investor Service. URL: https://www.moodys.com/sites/products/DefaultResearch/2006200000425644.pdf
18. Saitoh, F. (2016). Predictive modeling of corporate credit ratings using a semi-supervised random forest regression. IEEE International Conference on Industrial Engineering and Engineering Management, 429-433.
19. Staples, D., et al. (2013). Credit considerations for corporates in emerging markets. Moody's Investor Services. URL: http://www.moodys.com/research/Credit-Considerations-for-Corporates-in-Emerging-Markets--PBC_153812
20. Tharwat, A. et al. (2017). Linear discriminant analysis: A detailed tutorial. AI Communications, 30, 169-190.
21. 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.

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

Table 6. Rating scale of modelled 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[1], %	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 Investor Services (Moody's 2017)

1- Average one-year default rate calculated in 1983-2017 by Moody's Investor Service

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**	**EBITmargin**	**Eie**	**Dbc**	**RCF_d**	**Revenue_log**	**LogGDPpc**
Infl	1.00							
Exp	0.25	1.00						
EBITmargin	-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

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

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 of primary importance for researchers and practitioners as a lot of BRICS' ICs lack public credit ratings (CRs) from reputable rating agencies such as Moody's, Standard and Poor's or Fitch. The paper is aimed at filling the gap in the existing research as only very few efforts were focused on prediction of PCRs of ICs from entire BRICS IC community. The modelled variables are CRs of 208 BRICS' industrial companies assigned by Moody's at the year-end from 2006 till 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 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 modelling, Industrial company, Linear discriminant analysis, Ordered logit model, Artificial intelligence methods

JEL classification: C50-C53, G31-G33

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).

Moreover, large number of IC's debt in BRICS remained uncovered with the public credit ratings (PCRs) from international credit rating agencies (ICRA) such as Moody's, Standard and Poor's (S&P) or Fitch Ratings (Fitch). As a result, investors lack the publically available and reliable benchmarks to assess the IC's credit risk and need other reliable benchmark to replicate the missed PCRs. One of the widespread methods for building such benchmark is to model internal credit rating (ICR) which replicates the PCR. This underpins the relevance of this paper.

The motivation of the paper is underpinned by the narrow research and gaps in the available studies in the field of ICRs modelling. Only few efforts were based on instances from BRICS especially in field of industrial companies and were often limited to separate countries. Other gaps include employment of credit risk factors in models and application of methods of financial ratio computation techniques which did not match those in ICRA's methodologies. This casts doubts that constructed ICRs entirely replicate the PCRs. From practical standpoint, the existing studies did not compare the ability of different models to accurately replicate the PCRs. However, the practitioners strive for such comparison given that a lot of new methods including artificial intelligence and predictive analytics has become common due to advancing of computers.

The scientific novelty of this paper is the suggestion a new, more ambitious modelling approach to predicting PCRs which addresses the drawbacks mentioned above. Consequently, it makes analysis of constructed ICRs and their counterfactual testing more transparent. The distinctive features of approach are (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); (3) the application of the most actual data (for the period from 2006 to 2016) to control for the data stationarity. The practical novelty of the paper is providing comparison of forecasting power of wide range of statistical models and thus assisting debt professional to construct the ICRs which meet their goals, inputs and resources.

The outline is as follows. Section two represents the literature review in field of ICR modelling approaches and the analysis of drawbacks in existing studies. Section three presents the modelling approach and explores the data, the set of explanatory variables and the methods of modelling. Finally, in section 4, the accuracy of each method of reproducing Moody's PCRs is discussed and conclusions are formulated.

Internal credit rating modelling approaches: the literature review

The well-proven method of modelling 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). The choice of this method is underpinned by the fact that the PCRs demonstrated their ability to effectively discriminate between defaulters and non-defaulters (including ICs from BRICS) while reflecting more permanent changes in credit risks (Karminsky 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 freely published by ICRAs. Evaluation of assets' PDs are of paramount importance for practitioners. These probabilities are applied, for example, in establishing lending limits, calculation of regulatory capital in accordance to Basel III Accord requirements and estimation of expected credit losses under IFRS9 (Jarrow 2009, Basel 2001, Min 2017). Because of requirements of various regulations to apply PDs in assessing credit risk in banking and government finance, the substantial number of researches is focused on financial institutions (see Cucinelli et. al. 2018, Karminsky and Khromova 2016; Karminsky and Kostrov 2014) or municipal credit ratings (Ha'jek 2011) while we found a smaller number of efforts related to credit risk modelling of ICs.

The models applied for ICs included univariate studies, statistical methods and 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 and Giacomino 2007).

The results of studies of statistical methods demonstrated that OLR and OPR often outperformed other models such as linear or multiple discriminant analysis (Kamstra et. al. 2001). For example, Karminsky (Karminsky 2015) re-constructed the PCRs assigned by S&P and Moody's with the OPR. His sample included 215 ICs from 39 countries observed in 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 modelling. Depending on the various set of predictors, the model demonstrated the accuracy in range of 37%-43%. However, OLR and OPR possess the important limitation: multivariate normality assumptions for independent variables. These are frequently violated in financial data sets especially from BRICS. The other limitations of this study were (1) application of only one statistical method (OPR); and (2) limited timespan for modelling. Conversely, AI methods allow learning the particular structure of the model from the data. AIs usually outperforms statistical methods by forecasting power however they are complicated and hard to explain (Ha'jek and Olej 2011).

However, we identified several gaps in the existing efforts which motivate us to expand research in this area. Firstly, only limited number of efforts were focused on ICs from BRICS and all the studies are concentrated on separate BRICS countries but not on entire BRICS. The majority of studies were concentrated on USA or other developed economies such as that of Korea or Taiwan (Kumar and Bhattacharya 2006, Zan et al. 2004). However, research, indicated the difference in rating-driving factors between developed economies and BRICS (Bhogal 2017, Jiang and Packer 2017). Among the few efforts devoted to BRICS, Demeshev (Demeshev and Tikhonova 2014) compared the ability of several statistical and AI models (linear, quadratic and mixture DA, OLR, OPR and RF) to predict default of small and medium Russian ICs. He showed that statistical methods had less prediction power than that of AI, from which RF demonstrated the highest accuracy. However, Demeshev did not expand his study to large firms and/or ICs from other BRICS jurisdictions.

Secondly, only limited number of studies compared the predictive power of different statistical and AI methods. The vast majority of studies discussed only one or few models. For example, in his pivotal works, Altman (Altman et. al. 1977; Altman et. al. 2017; Altman et al. 2018) used only one statistical model – MDA. Metz and Cantor (Metz and Cantor 2006) worked out the UM model that converted ICs' financial metrics to implied ratings, took an appropriate weighted average of them and forecasted the Moody's PCRs. However, the comparison of forecasting power of various modelling approaches is demanded by practitioners to find the model which best fit their purposes.

Thirdly, many efforts applied limited (if any) non-financial factors in ICR models. For example, Kumar and Bhattacharya modelled ICRs with using only financial coefficients of issuers (Kumar and Bhattacharya 2006). The authors applied LDA and NN and used only financial variables in the modelling. The study confirmed that NN had a higher accuracy (79%) in comparison to LDA (33%). However, the rating process usually involve the extensive analysis of such corporate factors as macroeconomic conditions in the country of operations, the maturity of technologies, corporate governance and risk management practices, execution risks, etc. For example, Zhong (Zhong and Liu 2017) proved that political uncertainty directly influenced the credit risk. These factors need to be incorporated in the ICR models.

Fourthly, only limited number of studies analyzed the variety of artificial intelligence methods even though the latter have been gaining popularity among practitioners. For example, the scope of AI models in ((Ha'jek and Olej 2011) is limited to feedforward networks. Among the efforts which compared several AI models, Zan et al. re-constructed PCRs of ICs from Taiwan and the USA with AI models. (Zan et al. 2004). The SVM and NN demonstrated the slightly higher accuracy than OPR. However, this study has important limitations: (1) the application of abridged rating scale (rating classes only); and (2) the usage of the small sample.

Most important, in majority of research the explanatory variables describing credit risk factors or the way of their estimation significantly varied from those used in ICRA's methodologies. For instance, (Bellovary and Giacomino 2007) indicated that several financial metrics such as earnings before interest and taxes (EBIT) / interest, total debt/net worth or retained cash flow (RCF)/debt rarely applied in research. However, these metrics are cornerstone of the majority ICRA's methodologies of industrial companies (Moody's 2018). Additionally, the way of computation of financial variables in the majority of sources (see (Karminsky and Peresetsky 2007) or (Demeshev and Tikhonova 2016)) significantly varies from that of ICRAs (Moody's 2018). These cast doubts that constructed ICRs entirely replicate the original PCRs.

The conclusion is that most of the efforts described above were focused on developed markets (mainly US) while modelling of credit ratings is more demanded for BRICS corporates. Very few studies provided comparison of forecasting power of different models. However, such analysis is very important for ICs from BRICS because a lot of advantages of certain models reduce while limitations increase due to inefficiencies and structural problems in capital markets in these countries (Staples et al. 2013). The analyzed papers also had limitations such as small sample sizes, the limited set of models and/or explanatory variables used. We also observed differences in financial metrics computation methods between those in research and those in ICRAs' methodologies. This paper is aimed on filling these gaps.

The data, the variables and the methods

The motivation for modelling ICRs of industrial companies from BRICS

We modelled ICRs for industrial companies from BRICS[1]. This choice is underpinned by (1) growing systemic importance of BIRCS; (2) the increasing share of industrial production in these countries; and (3) the raising investors interest in ICs from BRICS (Staples 2013). World Bank statistics indicates: annual GDP growth in BRICS in 2008-2017 was 5.4% which was several times higher than those in developed world (0.8%) or other emerging (1.1%) or developing economies (2.6%). In the last decade the share of BRICS in global GDP has increased to 30% from 21.9%. GDP in BRICS in 2018-2022 will continue to increase by 4.7%, the higher pace that that in developed economies (1.5%) or other emerging (2.8%) and developing countries (2.5%)[2]. The projected investment spending in BRICS in 2017 was 33.2% vs. 18% in developed world and it will remain around 30% until 2030[3].

On the other hand, the significant share of industrial companies from BRICS lack PCRs from ICRA (Ratha et al. 2010). This is underpinned by (1) still developing financial infrastructure in these countries; and (2) regulatory restrictions on ICRAs operations (e.g. in Russia). The lack of public ratings increases the interest in reproducing PCRs by modelling of ICRs. As we showed in previous section, these are numerous research devoted to modelling of ICRs of financial institutions in BRICS (Karminsky and Khromova 2016, Karminsy and Kostrov 2014). However, the research devoted to modelling of ICRs of ICs is scare.

The data and the model's setting

For ICRs modelling we applied the mechanism developed in (Grishunin and Suloeva 2016). Its uses rating methodologies of Moody's as the framework. Our data set

---

[1] BRICS - is the acronym for an association of five major emerging national economies: Brazil, Russia, India, China and South Africa. These countries constitute over 40% of global population

[2] BRICS 2017. The role of BRICS in the world economy and international development. New Development Bank, https://reddytoread.files.wordpress.com/2017/09/brics-2017.pdf

[3] ibid

included 221 IC which 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).

The dependent variables in the model are the internal credit ratings which are converted by linear interpolation into numeric scores (Table 1). We applied Moody's rating scale for ICR system with alpha scores from Ca-Ba3 to Aa3-Aaa (see Table in Appendix). This scale consists of 13 grades, each grade is also mapped to a numerical scale from 1 to 13 and to idealized default probabilities. This approach is utilized by many ICRAs, for example, Moody's (Moody's 2018). Therefore, ICR is ordered, ordinal dependent variable which can take the values from 1 to 13.

Table 1. Rating scale of modelled 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[4], %	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 Investor Services (Moody's 2017)

The explanatory (independent) 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 cash flow debt coverage and the interest coverage. The remaining two are evaluated by subjective analysis of companies'

[4] Average one-year default rate calculated in 1983-2017 by Moody's Investor Service

business environment. For all components, we selected EVs which were the best match Moody’s methodologies (Moody’s 2018) (Table 2).

Table 2. List of financial explanatory variables

EV's description and notation	Formula and explanation
Revenue (Revenue), \$ million	IC's 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 which 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 $Dbc = \frac{Debt}{Equity}$
Cash flow debt coverage (RCF_d), %	$RCF\_d = \frac{OCF - CWC - Dividend}{Debt}$ OCF – operating cash flow of IC CWC – change in working capital

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 3.

Table 3. List of macroeconomic and dummy variables

EVs description and notations	Formula and explanations
Dummy variables	
IC is located in Russia	ICRAs usually apply 2-3 grids notching from the score obtained from agencies' rating grid for ICs domiciled in Russia. This is explained by developing instructional and financial infrastructure in the country (Moody's 2018). 1 – if IC is Russia based; 0 – if opposite
IC is located in China (China)[5]	Given the high sovereign rating of China (A1 stable), ICRAs do not apply notching for from

[5] ICRAs may also apply notching down from the score obtained from agencies' rating grid for ICs domiciled in Brazil, India or South Africa. This is explained by developing instructional and financial infrastructure in the country (Moody's 2018). ICs location in these countries was selected as a base category.

	the score obtained from agencies' rating grid for China-based ICs (Moody's 2018). 1 – if IC is China-based, 0 – if opposite
IC is owned by the government (Rtg)	1 – if IC is a government owned or significantly controlled entity, 0 – if opposite
IC is operating in a particular industry[6]: • 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)

The descriptive statistics of explanatory variables is presented in the Table 4.

Table 4. Descriptive statistics of independent 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
Interest coverage, x	Eie	8.2	228.1	-4.2	19.6
Gearing ratio, x	Dbc	0/48	1.47	0.02	0.20
Cash flow debt coverage, %	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

Statistical and AI methods and the modelling process

Linear discriminant analysis (LDA)

[6] Mining industry was taken as base industry.

LDA allows to discriminate between two or more groups of objects by multiple variables at the same time. The goal is the discovering such 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

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, $\boldsymbol{k = 1, 2, \ldots, p}$.

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 Polozov 2016):

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

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

Support Vector Machine (SVM)

SVM uses a linear model to implement non-liner class boundaries through nonlinear mapping 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 can be found the maximum margin hyperplane (MMH) which gives the maximum segregation between decision classes. The training example which are lying 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 which require classification of object to multiple grades (as required for ICR modelling). To account for non-linearity and multiple grades, the variable space is extended with the special kernel function. This allows to build the models with usage of separating hyperplane of various form (Ha´jek 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 applied γ=0.5.

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 modelled 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.

We trained our NN with the function neuralnet 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 which 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 the output units. To reduce the errors, the weights connected to the output units are adjusted to reduce the errors (a gradient descent). The network adjusts its weights incrementally until the NN stabilizes.

Random Forest (RF)

RF 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 performs bootstrap aggregation of a set of a decision trees. 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 repetitions 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). The studies showed that its advantages high prediction accuracy, avoidance of over-fit and robustness against high dimensional data (Saitoh 2016).

In the modelling 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 and 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 o 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.

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

Measuring the accuracy of ICR model

The accuracy of the j-th model can be measured with the modelling error - the difference between the actual PCR and modelled ICR.

$\Delta_{ij} = PCR_i - ICR_{ij}$

The accuracy of the model can be also characterized by the Type I and Type II errors. Type I errors are overstatement of modelled ICRs in comparison to actual PCRs. Type II errors are reverse – understatement of modelled 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 and Giacomino 2007). Therefore, the model which result in less Type I errors in relation to Type II error among the alternative models to be considered the best.

The error of each model, though, can be evaluated by the accuracy ratio which calculates the number of frequencies of each modelling errors. The accuracy ratio can be expressed by the following formula:

$$AR_{\Delta=z} = \frac{\sum_{i=1}^{M} w_{\delta=\Delta.i}}{M}$$

z- the condition, expressing the target Δ (e.g. z=0, z=1, z=2, $|z| \leq 1$, $|z| \leq 2$)

$w_{\delta=\Delta,j}$ – the binary variable, equals 1 if the modelling error Δi satisfies the condition of z, 0 if opposite.

i – the number of the observation in the validation sample

M – the total number of observations in the validation sample

The ICR model will have a good quality if the modeling errors not exceed two notches (Karminsky and Khromova 2016). In this case, the forecasted ICR will likely to remain within the same rating class (Aaa, Aa, Baa, Ba, B, Caa, Ca) as the actual PCR. Therefore, we chose for analysis the following modelling errors: $\Delta=0$ (no error), $\Delta=1$, $\Delta=-1$, $\Delta=2$, $\Delta=-2$, $|\Delta|\leq 1$, $|\Delta|\leq 2$.

Results

The modelling results

The results of ICR modelling with LDA are in Table 5 and Figure 1.

Table 5. Proportions of trace of each discriminant functions

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

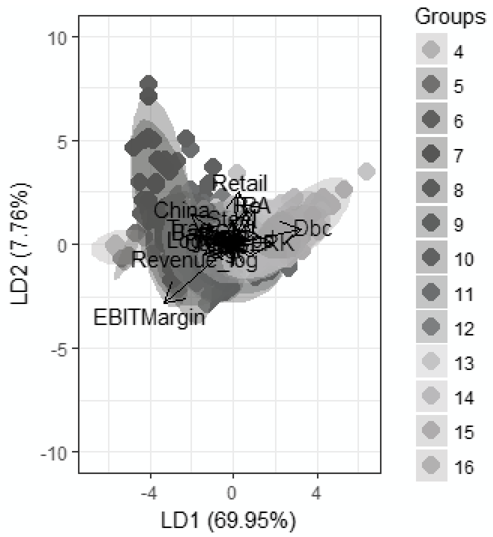

Fig. 1. Visualization of results of classification with LD1 and LD2

Application of the OLR gave the following results (Table 6).

Table 6. Result of ICR modelling 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**		

Note: ***, **, * - the coefficient is significant at levels of 1, 5, 10% respectively

Before ICR modelling by OLR we checked it for possible multicollinearity. Inter-factor correlation of the financial EVs is reported in Table 7.

Table 7. Matrix of inter-factor correlation for financial EVs in ICR model

	Infl	**Exp**	**EBIT margin**	**Eie**	**Dbc**	**RCF_d**	**Revenue_log**	**LogGDPpc**
Infl	1.00							
Exp	0.25	1.00						
EBITmargin	-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

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 2015) – 0.399 or lower.

The architecture of NN is presented at the Fig 2.

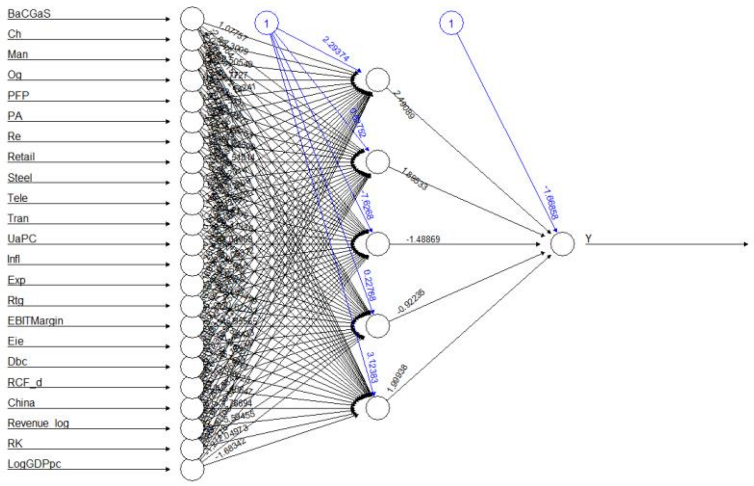

Fig. 2. The visualization of backpropagation NN used for ICR modelling

Lastly, the visualization of the most significant EVs in RF model are presented in Fig 3.

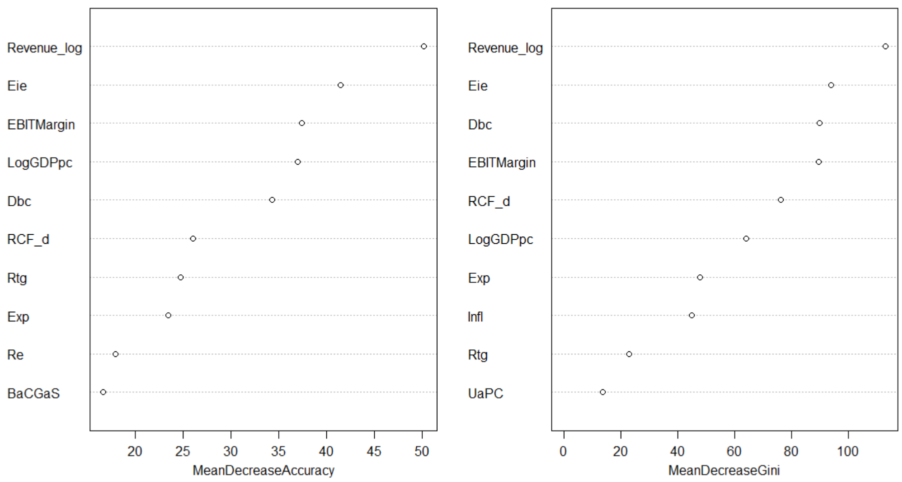

Fig. 3. The visualization of the most significant EVs in RF method

The summary of modelling outcome

The outcome of ICR modelling with above-mentioned statistical methods and its comparison to actual PCRs is presented in the Table 8. Negative z represents Type I error while positive z gives Type II error.

Table 8. The outcome of ICR modelling and its comparison to PCRs of BRICS' industrial companies

Model	Sample	Accuracy ratios ($AR_{Z=\Delta,}$), %						
		z=-2	z=-1	z=0	z=1	z=2	$\lvert z \rvert \leq 1$	$\lvert z \rvert \leq 2$
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

Our findings are summarized as follows. Artificial intelligence (AI) methods (SVM, NN and RF) outperform statistical modelling methods (LDA and OLR) by predictive power. This finding coincides with those in the existing research which made comparison between certain statistical and AI methods (Ha'jek and Olej 2011, Karminisky 2015). 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. AI methods also outperform statistical methods 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).

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 error, the number of Type I errors in AI model outcomes is very small and do not exceed 2.5% in total. Therefore, application of AI for modelling solves the problem of symmetrical errors mentioned in many research (Karminsky 2015, Karminsky and Khromova 2016).

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 2015, Grishunin and Suloeva, 2016). However, the accuracy of RF-based model deteriorates materially (to 58% on validation sample from 100% on training sample). Additional research is necessary for tuning the algorithm to limit such deterioration

Among the statistical 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 2015, Karminsky and Peresetsky 2007) of 38.8%-41.9%. However, the quality of our models is better than that reported in (Metz and Cantor, 2006) with prediction power of 21.5% of ICR modelling with OPR.

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% cases the prediction error does not exceed 1 notch (for SVM and NN – in almost 99% cases). The high prediction power of RF agrees with the previous findings of modelling ratings of financial institutions (Demeshev and Tikhonova 2014). In this research, the predictive power of RF varied from 60% to 72% depending on the industry.

Statistical methods despite lower prediction power that that of AI, have important advantage over AI methods – they are interpreted models and the researchers or practitioners can see the impact of each risk-factor in the model on the final rating outcome. AI models are “black boxes” because they cannot provide easy interpretation which risk factors are the most significant. This feature may limit the practical application of these models.

As a result of the modeling and specification of our models with the inclusion of country dummy variables, the estimates show us that the level of rating agency methodologies is the same for the BRICS countries. However, it is worth noting that for Russian companies, the cash flow debt coverage (RCF/Debt) and liquidity (Current Ratio) EVs at the one percent (1%) and five percent (5%) significance levels were particularly relevant.

The obtained empirical estimates also show that in terms of the level of solvency, Russian industrial companies (IC) are higher than Indian industrial companies and companies from South Africa. However, when comparing the results, Russian industrial companies have a worse financial picture than Chinese industrial companies.

An interesting observation was that the assessments of Russian industrial companies, in terms of debt level, are comparable with the results of Brazilian companies. However, the production capacities of the two countries (Russia and Brazil) are different. There is a negative correlation between the level of the company's assets and the assigned credit rating, this is typical for all BRICS countries. The proven hypothesis about the effect of the growth of the interest margin and it’s effect on the credit rating of an industrial company was confirmed. The presence of quadratic dependence and inclusion in the model variables showed that there is a positive credit rating increase with a high level of interest margin, which is typical for various countries (China, India, South Africa and Brazil).

Research limitations and directions of future research

Our study has several limitations. Firstly, the sample included large mature companies with significant history of operations and which are traded in global capital markets. Therefore, the additional research is necessary to test if the models will behave the same for medium and small size companies and less mature companies. Secondly, we considered only industrial companies from BRICS which limit the sample. Future direction of research will be to expand the sample by (1) including emerging economies outside the BRICS group; and (2) to extend the sample with the companies of the service sector. The latter is underpinned by growing purchasing power and consumption level.

Results shows that we were not able to achieve the increase in the prediction power of statistical models (LDA and OLR). This can be explained by that both models results in fixed coefficient liner indexes of the underlying risk factors (EVs). However, as ICRAs state, the relative importance of the EVs in the real rating process may vary with the values of other metrics (Metz and Cantor 2006). For example, for ICs from countries with high sovereign ratings the leverage may be less critical or for a highly leveraged issuer, interest coverage may be the most critical single factor, while for very low leveraged issuer, it may not be. The direction for future research is the application of cluster analysis for PCR prediction and changing the model specification to capture of rating relativeness.

Conclusion

This paper is devoted to comparison of the ability of various statistical methods to reproduce PCRs of BRICS's industrial companies using publicly available information. The motivation of the study is underpinned by (1) the narrow research and gaps in the available studies in the field of ICR modelling; and (2) the lack of research comparing the ability of various statistical and AI models to accurately replicate 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 artificial intelligence 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 modelled 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 modelled ICRs predicted actual PCRs with the errors not exceeding 1 notch. Unlike that 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 industrial companies from BRICS countries.

References

1. Altman, E., Haldeman, R., Narayanan, P. Zeta analysis. A new model to identify bankruptcy (1977). Journal of Banking and Finance, 1, 29-54
2. Altman, E., Iwanicz-Drozdowska, M., Latinen, E., Suvas A. (2017) Financial distress prediction in an international context: a review and empirical analysis of Altman's Z-score model. Journal of International Financial Management and Accounting, 26(2), 131-171
3. Altman, E., Essentato, M., Sabato, G. (2018) Assessing the credit worthiness of Italian SMEs and mini-bond issuers. Global Finance Journal, 38, 125-147
4. Balios, D., Thomadakis, S., Tsipouri, L. (2016) Credit rating model development: An ordered analysis based on accounting data. Research in International Business and Finance, 28, 122-136
5. Bhogal P. (2017). Rethinking the relevance of existing credit rating agencies to BRICS. ORF Occasional paper, 113
6. Basel Committee on banking supervision)2001). Consultative Document: The Internal Ratings-Based Approach. Basel, Switzerland: Bank for International Settlements
7. 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.
8. Cross-sector: the performance of Moody's corporate debt ratings – 2017Q2 (2017). Moody's Investor Service. URL: https://www.moodys.com/researchdocumentcontentpage.aspx?docid=PBC_1084629
9. Cucinelli, D., Di Battista, M., Marchese, M., Nieri, L. (2018) Credit risk in European banks: The bright side of the internal ratings based approach. Journal of Banking and Finance, 93, 213-229
10. Demeshev, B., Tikhonova A. (2014). Default Prediction for Russian Companies: Intersectoral Comparison. Economic Journal of Higher School of Economics, 18(3), 359-386
11. Financial statement adjustments in the analysis of non-financial corporations (2016). Moody's Investor Service. URL: https://www.moodys.com/research/Financial-Statement-Adjustments-in-the-Analysis-of-Non-Financial-Corporations--PBC_1050255
12. Jiang, X., Packer, F. (2017) Credit ratings of domestic global agencies: what drives the differences in China and how are they priced. BIS Working Papers, 648
13. Grishunin, S., Suloeva, S.: Development of the credit risk assessment mechanism of investment projects in telecommunication (2016). In: Galinina O., Balandin S., Koucheryavy Y. (eds) NEW2AN/ruSMART 2017. LNCS, vol 9870, 752-765.

14. Ha'jek, P. (2011) Municipal credit rating modelling by neural networks. Decision Support Systems, 51(1), 108-118.
15. Ha´jek, 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.
16. Jarrow R. (2009). Credit risk models. Annual Review of Financial Economics, 1, 37-68.
17. Kamstra, M., Kennedy, P., Suan, T. (2001) Combining bond rating forecasts using logit. Financal Review, 37, 75–96
18. Karminsky, A. (2015). Corporate ratings models for emerging markets. Corporate Finance, 3, 19-29.
19. Karminsky, A., Khromova E. (2016). Modelling banks' credit ratings of international agencies. Eurasian Economic Review, 6, 341-363.
20. Karminsky, A., Kostrov, A. (2014). The probability of default in Russian banking. Eurasian Economic Review, 4(1), 81–98.
21. Karminsky, A., Peresetsky, A. (2007). Models of ratings of international rating agencies. Prikladnaia Ekonometrika (Applied Econometrics), 1, 3–19. (in Russian).
22. Karminsky, A., Polozov, A. (2016): Handbook of Ratings: Approaches to Ratings in the Economy, Sports, and Society, Springer, 338.
23. 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.
24. Lee, Y. C. (2007). Application of support vector machines to corporate credit rating prediction. Expert Systems with Applications, 33, 67–74.
25. Liu, J., Zhong, R. (2017). Political uncertainty and a firm's credit risk: evidence from the international CDS market. Journal of Financial Stability, 53-66
26. List of rating methodologies (2018). Moody's Investor Service. URL: https://www.moodys.com/research/List-of-Rating-Methodologies--PBC_127479
27. Metz, A., Cantor, R. (2006). Moody's Credit Rating Prediction Model. Special Comment. Moody's Investor Service. URL: https://www.moodys.com/sites/products/DefaultResearch/2006200000425644.pdf
28. Miu, P., Ozdemir B. (2017) Adapting the Basel II advanced internal-ratings-based models for International Financial Reporting Standard 9. Journal of Credit Risk, 13(2), 30-62.
29. Ratha, D.; De, P.K., Mohapatra, S. (2010). Shadow sovereign ratings for unrated developing countries. World bank working papers. 2010. P. 2.
30. Saitoh, F. (2016). Predictive modeling of corporate credit ratings using a semi-supervised random forest regression. IEEE International Conference on Industrial Engineering and Engineering Management, 429-433.
31. Staples, D., et al. (2013). Credit considerations for corporates in emerging markets. Moody's Investor Services. URL: http://www.moodys.com/research/Credit-Considerations-for-Corporates-in-Emerging-Markets--PBC_153812

32. Tharwat, A. et al. (2017). Linear discriminant analysis: A detailed tutorial. AI Communications, 30, 169-190.
33. 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.