

Two-step classification method based on genetic algorithm for bankruptcy forecasting 破产



Yuri Zelenkov^a, Elena Fedorova^{a,*}, Dmitry Chekrizov^b

^a National Research University Higher School of Economics, 101000, 20 Myasnikitskaya str., Moscow, Russia

^b "Globalstar – SpaceTelecommunications" Joint Stock Company, 117485, 7 Butelerova str., Moscow, Russia

ARTICLE INFO

Article history:

Received 9 November 2016

Revised 16 June 2017

Accepted 15 July 2017

Available online 17 July 2017

Keywords:

Bankruptcy

Bankruptcy forecasting models

Ensembles of classifiers

Features selection

Genetic algorithm

ABSTRACT

By present, many models of bankruptcy forecasting have been developed, but this area remains a field of research activity; little is known about the practical application of existing models. In our opinion, this is because the use of existing models is limited by the conditions in which they are developed. Another question concerns the factors that can be significant for forecasting. Many authors suggest that indicators of the external environment, corporate governance as well as firm size contain important information; on the other hand, the large number of factors does not necessary increase predictive ability of a model. In this paper, we suggest the genetic algorithm based two-step classification method (TSCM) that allows both selecting the relevant factors and adapting the model itself to application. Classifiers of various models are trained at the first step and combined into the voting ensemble at the second step. The combination of random sampling and feature selection techniques were used to ensure the necessary diversity level of classifiers at the first step. The genetic algorithms are applied at the step of features selection and then at the step of weights determination in ensemble. The characteristics of the proposed method have been tested on the balanced set of data. It included 912 observations of Russian companies (456 bankrupts and 456 successful) and 55 features (financial ratios and macro/micro business environment factors). The proposed method has shown the best accuracy (0.934) value among tested models. It has also shown the most balanced precision-recall ratio. It found bankrupts (recall = 0.953) and not bankrupts (precision = 0.910) rather accurately than other tested models. The ability of method to select the task-relevant features has been also tested. Excluding the features that are significant for less than 50% of the classifiers in the ensemble improved the all performance metrics (accuracy = 0.951, precision = 0.932, recall = 0.965). So, the proposed method allows to improve the advantages and alleviate the weaknesses inherent in ordinary classifiers, enabling the business decisions support with a higher reliability.

© 2017 Elsevier Ltd. All rights reserved.

精确的召回率

多样

缓和

1. Introduction

The problem of bankruptcy forecasting is one of the most actively studied nowadays among practical and theoretical issues of company management. Assessment of the current financial status and determination of bankruptcy probability are of interest to shareholders, suppliers, creditors and others aiming to deal with perspective and reliable business partners.

Although more than 150 bankruptcy models employing above 750 different factors have been proposed, these models are rarely used in practice (Bellovary, Giacomino, & Akers, 2007). In our opinion, this is since different economic environments have various properties that do not allow reusing models and related sets

of factors in other conditions. This fact has been confirmed by comparative studies of models for different countries (Liang, Tsai, & Wu, 2015; Peng, Wang, Kou, & Shi, 2011). For developing countries, with the economic structure significantly different from the developed countries this fact is particularly significant (Fedorova, Gilenko, & Dovzhenko, 2013). In addition to the financial ratios, new factors may be required for forecasting in such conditions (Liang, Lu, Tsai, & Shih, 2016; Tinoco & Wilson, 2013). However, it is not possible to predict in advance which indicators will affect the predictive ability, so the features selection is necessary.

In this paper, we design a predictive model that can be easily adapted to any new situation, both in terms of model architecture and the features used. We combine the following techniques: at the first stage, we select the significant features for ordinary classifiers, which are combined into a voting ensemble in the second stage. For both stages (features selection and determination of the weights of the ensemble) a genetic algorithm is used.

* Corresponding author.

E-mail addresses: yuri.zelenkov@gmail.com (Y. Zelenkov), ecolena@mail.ru (E. Fedorova), chekrizovdv@mail.ru (D. Chekrizov).

The proposed method has been tested on the balanced set of data that consists of 912 observations of Russian companies (456 bankrupts and 456 successful companies) and 55 features (financial ratios, macro and micro business environment factors). The research data set was divided into a training set and a test set in a proportion 80/20. The proposed method has shown the best accuracy (0.934) among tested models on the test set. It has also shown the most balanced precision-recall ratio. It determined bankrupts (recall=0.953) and not bankrupts (precision=0.910) with a higher accuracy compared to the other tested models.

Another advantage of the proposed method is the ability to select the **conditions-dependent features**. Elimination of features significant for less than 50% of classifiers in the ensemble increases all the performance metrics (accuracy=0.951, precision=0.932 and recall=0.965). **条件依赖特征 消除**

2. Related works

Nowadays a number of studies are devoted to the problem of bankruptcy probability assessment. All researches of this subject can be conditionally divided into two groups. The first one consists of the researches **concerning the choice of measurable characteristics (features) set providing a high precision forecasting** (Kumar & Ravi, 2007). The most extensive of these studies concerns the indicators of financial condition (du Jardin, 2016; Fedorova et al., 2013; Lee & Yeh, 2004; Zięba, Tomczak, & Tomczak, 2016). There are works which are focused on corporate management factors (Bredart, 2014; Chan, Chou, Lin, & Liu, 2016; Lee & Yeh, 2004; Liang et al., 2016; Salloum & Azoury, 2012), environmental factors (Alifiah, 2014; Bruneau, de Bandt, & El Amri, 2012; Delas, Nosova, & Yafinovich, 2015; Duffie, Saita, & Wang, 2007; Karas & Režňáková, 2014; Tinoco & Wilson, 2013; Vlamis, 2007), level of legislation development (Rowoldt & Starke, 2016), etc. In most cases, the forecasting model of bankruptcy includes characteristics from all of these factors sets (Chan, et al., 2016; Liang et al., 2016; Tinoco & Wilson, 2013). According to Bellovary et al. (2007), who reviewed bankruptcy prediction studies from 1930 to mid-2000s, a total of 752 different factors are used in researches, yet 674 of the factors are utilized in only one or two studies. The number of factors considered in one study ranges from 1 to 57 and the average number of factors is 10.

The second group of studies focuses on forecasting methodology. The bankruptcies prediction is commonly considered as a two-class classification problem: there is a set of same type objects (in our case – firms) which belong to known classes (in our case – bankrupts and non-bankrupts); the goal is to design an algorithm capable of classifying analogical objects of unknown class.

Aziz and Dar (2006) stated that bankruptcy prediction models can be divided into three main categories: statistical models, artificial intelligence expert models and theoretical models. According to Bellovary et al. (2007), prior to the 1990s, the basic prediction methodologies were based on statistical methods including logit and probit models, and the most common was **discriminant analysis (DA)**, first applied in the seminal work of Altman (1968). DA attempts to derive a linear combination of features, which best 'discriminate' between the classes. Altman (1968) determined 5 financial ratios, which linear combination helps to calculate a discriminant score for the US manufacturing firms; this model is well-known as Z-score.

After the 1990s, artificial intelligence expert systems including machine-learning techniques became the primary method for bankruptcy prediction. Most often artificial neural networks (ANN) were used (Bellovary et al., 2007), especially the multilayer perceptron (MLP). It should be noted, however, that many researchers stated that ANNs have some shortcomings (Dos Gordini, 2014; Santos, E.M., & Maupin, 2009): (1) a researcher should have great

experience in order to select the control parameters properly, (2) it is difficult to generalize the results due to overfitting, and (3) ANNs have lack of explanatory power, so it is difficult to explain the prediction results.

In addition, some authors have investigated the applicability of other artificial intelligence methods to bankruptcy prediction, for example, the case-based reasoning (Ahn & Kim, 2009; Sartori, Mazzucchelli, & Di Gregorio, 2016).

Bellovary et al. (2007) come up with the following conclusions in their review: (1) there are over 150 models available, many of which have shown high predictive ability, so the focus of future research should be on using the existing models as opposed to the development of new ones; (2) a large number of factors does not necessary increase model's predictive ability; (3) existing models are poorly used in practice, so researchers should attempt to establish stronger connection with practice.

We disagree with the first limitation. First, the existing gap with practice, noted by Bellovary et al. (2007), is the consequence of the bounded applicability of available models. Many studies show that it is impossible to use models prepared for a specific industry, market or country in other conditions. Liang et al. (2015) examined data of four different countries (Germany, Australia, Taiwan, and China). Their results indicate that there is no single best combination of feature selection method and the classification technique over the four data sets. Similar results were obtained by Peng et al. (2011), who analyzed bankruptcies in Japan and Korea. Fedorova et al. (2013) showed that the Altman Z-score, based on the five financial ratios identified by Altman (1968), does not yield acceptable results for Russian companies. Moreover, an attempt to build a new DA model that separates Russian bankrupts and successful firms also does not provide significant results (the confirmation of this statement is presented below in a section describing research results). In addition, the size of the firms studied significantly affects the model. Gordini (2014) showed that the models for small and medium-sized enterprises differ from those for large enterprises. Therefore, the development of new models (or, to be precise, the adaptation and combination of known techniques) and use of new factors are required for each specific case.

At the same time, we believe there is no need in limiting ourselves to using models proposed 20 years ago and earlier. Many new machine learning techniques are known today, which can also be successfully applied to the problem of bankruptcy prediction that is confirmed by flow of scientific publications. Both ordinary methods of classification (logistic regression (LR), k nearest neighbors (kNN), support vector machine (SVM), naive bayes (NB), decision trees (DT)) and methods based on classifiers ensembles are used to solve this problem. Numerous works have been devoted to comparative efficiency analysis of various classification techniques applied to bankruptcy forecasting in different countries (Chang & Yeh, 2012; Fedorova et al., 2013; Heo & Yang, 2014; Liang et al., 2016; Peng et al., 2011; Tsai, Hsu, & Yen, 2014; Zhou, Lai, & Yen, 2014).

Secondly, the use of new factors in addition to financial ratios can provide new useful information. We agree that the large number of factors in general does not increase the accuracy of the model, and it is not possible in advance to predict which of them will be informative. Therefore, it is advisable to combine an increase in the number of factors with methods of features selection (Liang et al., 2016).

Thus, the goal of our work can be formulated as following: to develop a more or less universal method that allows to build an effective classifier, which can adapt to different conditions. The proposed method should (1) select significant features and (2) determine the architecture of the classifier. To achieve this goal, we propose using evolutionary techniques, namely genetic algorithm (GA). Such approach should ensure high applicability in practice.

Some researchers use evolutionary techniques for parameters optimization in addition to more or less determined standard methods of classifiers creation. In this case, the optimal combination of the classifier parameters is searched in the space of their possible values, the optimization goal (i.e. fitness function) is defined usually as a minimal error of classification. For example, [Gordini \(2014\)](#) used GA to build a system of logical rules based on the sign and the cutoff value of the selected financial ratios for Italian small and medium enterprises. This helped the author to create discriminative system that outperforms the results of LR and SVM classifiers, especially in reducing the Type II error. Similar approach was used by [Gorczalczy and Rudzinski \(2015\)](#), who proposed an automatic design of fuzzy rule-based classifiers using multi-objective evolutionary optimization algorithms. The genetic algorithm is also often employed to select significant features ([Acosta-Gonzalez & Fernandez-Rodriguez; 2014](#)), sometimes this approach is combined with parameter optimization. For example, [Zhou et al. \(2014\)](#) used the genetic algorithm for simultaneous SVM's feature selection and parameter optimization. Such approach allows receiving the classifiers that consider particular features of subject domain more precisely. One of the elements required for accurate prediction by using an ensemble is recognized to be an error "diversity" ([Brown, Wyatt, Harris, & Yao, 2005](#)). If there are many different classifiers to be combined, one would expect an increase in the overall accuracy when combining them, if they are diverse. [Kim and Kang \(2012\)](#) proposed a genetic algorithms-based coverage optimization system for ensemble learning.

Let us note that the majority of ensembles creation techniques operate with the classifiers of one type. The alternative direction in the development of [classification methodology is the combination of several classifier systems](#), which can be built following either the same or different models and/or data sets building approaches ([Wozniak, Grana, & Corchado, 2014](#)). Such systems exceed the traditional approaches relying on classifiers of one type. They use a combination of various rules and methods of information extraction for decision making. The approach based on creation of ensemble from classifiers of various type is known as Multi-Classifiers Systems (MCS) and is actively developing.

This paper follows the MCS approach and describes a two-stage classification method (TSCM) using genetic algorithms. In our TSCM method we suggest using a combination of several approaches: training of individual classifiers at training set with simultaneous features selection for each individual classifier based on GA; determination of each classifier significance in ensemble also through GA.

The choice of ordinary classifiers is the first problem in case of algorithm creation. [Thammasiri and Meesad \(2012\)](#) used DT, SVM and ANN. [Tsai \(2014\)](#) used a combination of MLP, DT, and LR as the hybrid classifier, other models of classifiers haven't been researched. In our research, we choose the most popular unary classifiers kNN, LR, NB, DT and SVM.

The second problem concerns the feature selection. The feature selection problem deals with selection of an [optimal relevant set](#) of attributes that are necessary for the recognition process ([Han, Kamber, & Pei, 2012; Mukhopadhyay, Maulik, & Bandyopadhyay, 2014](#)). It helps reduce the dimensionality of the measurement space and [facilitates](#) the use of easily computable algorithms for efficient classification. [Liang et al. \(2015\)](#) noted that since there is no general agreement about financial ratios as input features for model development for financial distress prediction. A number of studies consider feature selection as a pre-processing step in data mining before modeling. Some of the collected variables may contain noise that could affect the prediction result. On the other hand, if too many features were used for data analysis, it can cause high dimensionality problems. The simplest algorithm of feature selection is to test each possible subset of features finding the

one which minimizes the error rate, and the choice of evaluation metric heavily influences such method ([Guyon & Elisseeff, 2003](#)), in addition, a full bust of all possible combinations requires a lot of time. **粒子群 模拟退火**

Subset selection approaches include a lot of techniques, significant part of which is also inspired by nature: genetic algorithm, [particle swarm, simulated annealing](#), etc. For example, [Acosta-Gonzalez and Fernandez-Rodriguez \(2014\)](#) used GA for selecting the financial ratios employed in the LR model. In our research, we will also rely on [GA as evolutionary technique allows to research parameters space with the smallest costs and to choose optimum feature set for each individual classifier](#).

Also, the random subspace method (RSM) offered by [Ho \(1998\)](#) is widely applied to feature selection, this approach is also called attribute bagging ([Bryll, 2003](#)) or feature bagging. According to this approach, the features are randomly sampled with replacement for each classifier. This causes individual classifiers to not over-focus on features that appear highly predictive/descriptive in the training set, but fail to be as predictive for points outside that set. When RSM is combined with bagging of decision trees, the resulting models are called random forests (RF). Similar approach is used also in models which are named extremely randomized trees ([Geurts, Ernst, & Wehenkel, 2006](#)) or ExtraTrees (ET). Peculiarity of this technique is that top-down splitting in the tree classifier is randomized instead of computing the locally optimal feature/split combination. [We also use random subspace method when training individual classifiers to avoid over-focusing](#).

The third problem is the choice of classifiers ensemble architecture. We use the voting classifier since it is the simplest way to realize MCS. [The advantage of voting classifier use over AdaBoost, bagging, random subspace](#), and majority voting has been proved in the research of [Onan, Korukoglu, and Bulut \(2016\)](#).

Thus, we solve a quite standard problem regarding the development of the model with high prognostic ability. The higher accuracy of our method (TSCM) is achieved through the inclusion of a stage in which a genetic algorithm-based feature selection is made for each individual classifier that is trained on the random subspace.

In addition, we have significantly expanded the set of variables for analysis. We added to the traditional financial ratios (liquidity, financial stability, turnover and profitability) the ratios used in the Altman model, and the factors chosen according to the frequency of applying in various studies ([Bellocary et al., 2007](#)). Similar sets were used in many works ([du Jardin, 2016; Fedorova et al., 2013; Lee & Choi, 2013](#)). Also, we consider that it is important to estimate the external environment factors of the firm. Researchers in a bankruptcy forecasting area often use various external indicators and most often economic conjuncture indicators, such as: GDP and its derivatives, various market indices, PMI (purchasing managers index), etc. ([Alifiah, 2014; Karas & Režňáková, 2014; Vlamis, 2007](#)). In our research, we adhere to the classification of external factors proposed by [Delas et al. \(2015\)](#). However, we have excluded some indirect indicators (for example, the cultural environment) and concretized macroeconomic external factors, attributing them to an economic environment or to a policy of the Central Bank (the rate of license revocation from credit institutions, the market indexes, the price index of manufactured goods). In addition, the characteristics of the firms being investigated were included in the data set (e.g. inclusion of the firm in the public list of unreliable suppliers).

3. Two step classification method

The main issues in the [混合 系统拓扑](#) [hybrid](#) classifier system (MCS) design ([Wozniak et al., 2014](#)) are: [system topology](#) (how to interconnect individual classifiers), [ensemble design](#) (how to drive the [合奏](#) [ensemble](#)).

最佳相关集

促进

eration and selection of a pool of valuable classifiers) and **fuser** 定义 design (how to build a decision combination function).

3.1. System design

We use a parallel architecture as the overwhelming majority of MCS reported in the literature is structured in a **parallel topology** 并行拓扑 (Kuncheva, 2004; Wozniak et al., 2014). In this architecture, each classifier is fed by the same input data, so that the final decision of the ensemble is based on the individual classifiers outputs, obtained independently.

3.2. Ensemble design

相互补充

参与 The design algorithm of hybrid ensemble should support the involvement of mutually complementary individual classifiers that provide high diversity and accuracy. Two problems should be considered on the way to design such algorithm. First, there is no commonly accepted methods of diversity measurement (Brown et al., 2005). In addition, according to Kuncheva (2004), it is impossible to raise diversity of classifier pool through manipulation with inputs and model of individual classifiers. Secondly, it is impossible to predict what classifiers can be complementary. It depends on type of task and specific data set.

Therefore, we suggest using the combination of three approaches: (1) training of individual classifiers on subsets of training data and the calculation of their accuracy as the average accuracy on these subsets (random sampling); (2) features selection for each individual classifier based on GA; (3) calculation of the weight of each classifier in ensemble also by GA.

3.3. Fuser design

We use voting ensemble with majority voting rule. Output value is formed as the weighted sum of individual classifiers responses. For this purpose, the class of each object is determined by values 1 (bankrupt) and -1 (not bankrupt). The sign of the ensemble outcome determines to the object class, the outcome absolute value corresponds to the degree of confidence.

3.4. Two-stage ensemble evolutionary training

The offered algorithm is implemented in two stages. At the first stage, training of individual classifiers and selection of the most adequate feature set is made for each of them. As it is mentioned above, it is impossible to predict what classifiers will have complementary properties, so one can use any available models in this phase. We use the most widely applied models of ordinary classifiers: kNN, NB, LR, DT and SVM.

3.4.1. First stage. Training of classifiers set

Coding. The features set used for the ordinary classifier training is coded by an array of bits $G(N)$, here N is number of features in the initial dataset. Array elements can take values 0 or 1. If the element of G is equal to 0, the corresponding feature is excluded from training set.

Initial population. Initial population of size s is created from classifiers of one type. Value 1 is assigned to all elements of a genotype G of each population member. Thereby training of each classifier begins with full range of features.

Selection. The best individual in population is always copied in new population without any changes (the principle of elitism). Selection of other individuals is rank-based. 精英

Mutation. Mutation is applied to randomly selected element of genotype G of selected individual, its value is replaced by opposite, i.e. 0 becomes 1, and 1 becomes 0.

Algorithm 1

```
for  $T_{train}, T_{test}$  in split( $T, n_{fold}$ )
   $C = \text{train}(T_{train})$ 
   $s[i] = \text{accuracy}(C, T_{test})$ 
 $f = \text{mean}(s)$ 
 $f.\text{std} = \text{std}(s)$ 
```

后代

Crossover. Two-points crossover is applied with probability p_c to two selected individuals, their offsprings are added in new population.

Fitness function. Fitness of each population member is calculated as an average classification accuracy value according to an Algorithm 1. The initial training set T contains L observations. It is randomly split in two not crossed subsets: T_{train} of size $(1 - 1/n_{fold})L$ and T_{tests} of size L/n_{fold} . The first set is used for classifier C training, the second one is used for its accuracy assessment. This procedure is repeated n_{fold} times, fitness is calculated as a mean accuracy by all n_{fold} attempts. Such approach allows to avoid overfitting. The accuracy calculated as

$$\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

Here TP is true positives (positives classified as positives), FP is false positives (negatives classified as positives), TN is true negatives (negatives classified as negatives) and FN is false negatives (positives classified as negatives).

At the first stage of training a set of classifiers types that will be used for ensemble creation is specified (denote the number of types in this set by M). The classifier of each type is trained according to the algorithm described above. The set of the trained classifiers is transferred to the following stage.

3.4.2. Second stage. Creation of ensemble

In the second stage, the voting ensemble with majority voting rule is designed from the set of the classifiers trained at the first stage.

Coding. The ensemble is coded by an array of w real numbers, big or equal to zero of size M . They set value of weight coefficient to corresponding classifier. During creation of initial population, the elements $w_i, i = 1 \dots M$ are set in a random way, at the same time the normality condition is satisfied $\sum_i w_i = 1$.

Selection. Selection rules are the same as at the first stage: principle of elitism and selection according to the rank of individual.

Mutation. Mutation is applied to all weight coefficients w of the selected individual. These weights w_i are changed by the random values v_i uniformly distributed in the interval $(-0.1; 0.1)$. If negative value of w_i is received, it is replaced by 0. At the same time, the normality condition isn't satisfied. These parameters were determined during experimental launches of an algorithm.

Crossover. Two-point crossover operation is similar to the crossover at the first stage.

Fitness function. Fitness of ensemble is calculated as the accuracy value on training set T . Let C_E be the class of object that is calculated as the weighed sum of outcomes of ordinary classifiers: $C_E = \sum_{i \in M} w_i C_i$, where C_i is an outcome of i th ordinary classifier. If the sign of C_E matches the sign of object class (1 for bankrupts and -1 for the successful company), then an object is considered as correctly recognized. Absolute value of C_E corresponds to degree of confidence.

The program code, which realized the method, was developed using the Python language on base of machine learning library scikit-learn (Pedregosa, et al., 2011).

Table 1
Results of basic classifiers training (the first stage of the method).

Classifier model	Training set						Test set			Weights in ensemble w_i
	Before training (N = 55)	After training					Accuracy	Precision	Recall	
	Accuracy	Accuracy	Accuracy Growth	Precision	Recall	Number of features (N)				
kNN	0.830	0.871 ± 0.055	1.05	0.851	0.911	34	0.831	0.787	0.871	0.283
LR	0.808	0.829 ± 0.024	1.03	0.810	0.868	51	0.825	0.773	0.882	0.013
NB	0.557	0.602 ± 0.070	1.08	0.598	0.669	33	0.579	0.540	0.635	0.122
DT	0.878	0.892 ± 0.028	1.02	0.895	0.892	46	0.847	0.852	0.812	0.184
SVM	0.848	0.855 ± 0.012	1.01	0.839	0.884	54	0.845	0.816	0.856	0.287

4. Analysis of data

To assess the quality of the proposed method a balanced data set was used. It consists of 912 observations (456 bankrupts and 456 successful companies) of Russian firms and 55 features. The original data in the data set were normalized so that their values lied in the range $(-1; 1)$. This data set was divided into 2 sets – the training set and the test set, taken in the proportion of 80/20, which included 729 and 183 observations, respectively.

The list of variables for analysis included standard financial ratios (liquidity, financial stability, turnover, and profitability), the ratios of Altman (1968) model and the financial ratios selected according frequency of their occurrence in researches in various countries (Bellovary et al., 2007). We also added indicators that reflect the economic situation and the policy of the Central Bank, and factors that describe firm's non-financial characteristics: the presence of the government control, the presence of economic sanctions of different States, market share, inclusion of the firm in the public list of unreliable suppliers, etc. Table 4 lists all features selected for method evaluation.

5. The results and discussion

The results of basic classifiers training (the first stage of a method) are given in Table 1, for all classifiers the following parameters were used: the population size is 50, number of generations is 50, $n_{fold} = 4$, $p_c = 0.5$. Accuracy before training was calculated before start of a genetic algorithm when the features set included all 55 elements. Such widely known models as kNN, LR, NB, DT and SVM were used as the basic classifiers with the default parameters of scikit-learn library. The average accuracy reached after training is given in the column with caption “after training” with borders of 95% confidence interval. The highest accuracy is received for DT method (0.892) with the border of 95% confidence interval ± 0.028 . The narrowest confidence interval belongs to SVM method (± 0.012).

The most essential indicators of the classification quality are the precision and the recall, which are defined by the following formulas:

$$precision = \frac{TP}{TP + FP}, \quad recall = \frac{TP}{TP + FN}.$$

The precision is the ability of the classifier not to label a sample as positive that is actually negative. The recall is the ability of the classifier to find all the positive samples. These values correspond to the two types of errors. Type I errors are the misclassification of bankrupt firms as non-bankrupts, this error is measured via recall. Type II errors are, on the contrary, non-bankrupt firms misclassified as bankrupts, this error is measured via precision.

Also, Table 1 shows the number of features, which were recognized as important in the process of training. Various basic classifiers choose different number of features, the SVM method selects the largest number (54), and the NB method selects the smallest number of features (33).

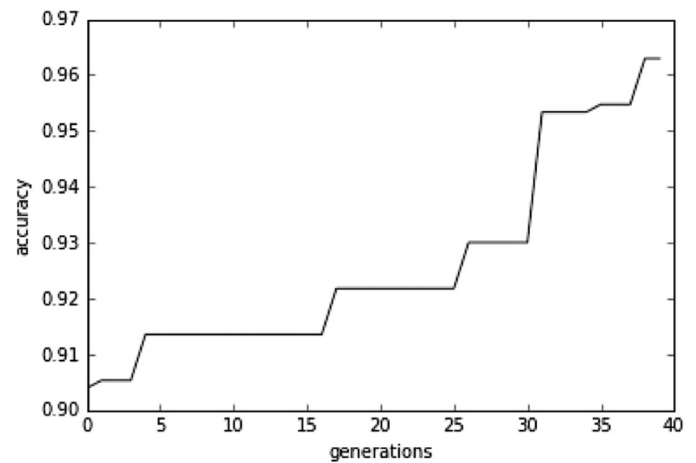


Fig. 1. Training of ensemble at training set. Classification accuracy rate vs number of generations.

The values of accuracy, precision and recall on the test sample are given in the columns with caption “Test set”. The DT model also demonstrates the highest accuracy on a test set.

Fig. 1 presents the evolution of accuracy during training of ensemble (the second stage of method), which combines classifiers trained at the first stage. The following parameters are used: the population size is 40, number of generations is 40, probability of crossover $p_c = 0.5$. As follows from Fig. 1, the final accuracy of ensemble is 0.963. Weights w_i of the trained ensemble are specified in an extreme right column of Table 1. The SVM model has the largest weight (0.287), the second is kNN model (0.283).

Table 2 shows the results of training and testing of the proposed method in comparison with the results of other popular methods. First, we tested the two most popular models of bankruptcy forecasting (Bellovary et al., 2007), namely DA and ANN. It can be seen from Table 2 that the Z-score with the coefficients determined in Altman (1968) has a significant Type II error, it misclassifies many successful firms as bankrupts (precision = 0.467, recall = 0.906). This confirms that it is impossible to reuse models prepared for certain industries, markets and countries in other conditions. To further investigate this problem, the new discriminant function was calculated with help of Linear Discriminant Analysis (LDA) on the full set of features. It significantly improved the results of classification (see Table 2, line ‘LDA’), but the analysis of the scaling vector for the new model showed that the largest contribution to class separation is provided not by the features identified by Altman (1968). In descending order of significance, the new top-5 factors are R5 (profitability), S8 (the ratio of own working capital), T3 (total assets turnover), E22 (growth of HSBC PMI index) and T4 (net income to net capital). A significantly better discriminant function can be obtained by a Quadratic Discriminate Analysis (QDA). We note, however, that the

Table 2
Accuracy of various models of bankruptcy prediction.

Method	Training set			Test set		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
POPULAR BANKRUPTCY PREDICTION MODELS						
Altman's Z-score	0.506	0.509	0.865	0.475	0.467	0.906
LDA	0.890	0.888	0.898	0.880	0.846	0.906
QDA	0.973	0.953	0.995	0.902	0.868	0.976
MLP (45)	0.986	0.976	0.997	0.923	0.875	0.976
MLP (40,5)	0.997	0.995	1.000	0.929	0.909	0.941
ENSEMBLE METHODS						
BAGGING – DT	0.999	0.997	0.999	0.913	0.906	0.906
BAGGING – KNN	0.962	0.939	0.989	0.923	0.874	0.976
BAGGING – SVM	0.909	0.902	0.922	0.918	0.880	0.953
BAGGING – NB	0.658	0.657	0.687	0.497	0.467	0.576
BAGGING – LR	0.896	0.889	0.908	0.880	0.846	0.906
ADABOOST – DT	0.999	1.000	0.999	0.880	0.846	0.906
ADABOOST – SVM	0.509	0.509	0.999	0.464	0.464	0.999
ADABOOST – NB	0.824	0.846	0.801	0.776	0.762	0.753
Gradient Boosting	0.984	0.969	0.999	0.858	0.824	0.882
Voting Classifier	0.929	0.916	0.946	0.891	0.857	0.918
RandomForest	0.999	0.999	0.999	0.831	0.855	0.765
ExtraTrees	0.999	0.999	0.999	0.929	0.901	0.953
PROPOSED METHOD						
TSCM	0.963	0.943	0.987	0.934	0.910	0.953

method of features selection based on GA proposed here (the first stage of TSCM) works very poorly for the LDA and QDA classifiers, tests have shown that there is no reduction in number of features. Therefore, classifiers based on the discriminant analysis were not included in the list of basic classifiers of TSCM.

Of the many ANN models, we tested Multi-Layer Perceptron (MLP), as it is the most popular model in the field of bankruptcy prediction. After the tests, two models were selected: a network with one hidden layer with 45 neurons (MLP (45)) and a network with two hidden layers consisting respectively of 40 and 5 neurons (MLP (40,5)). Both MLP models showed results that exceed the results of discriminant analysis, and this is consistent with Bellovary et al. (2007) and Liang et al. (2016). However, our tests showed that for MLP models, the feature selection also shows a relatively poor performance. In addition, the training time of these models significantly exceeds that of the other models. Therefore, MLP models were also not included in the set of basic classifiers.

A significant number of ensemble methods have already been developed. Results of their comparison on training and test sets described above are also shown in Table 2. We tested five bagging methods (with DT, kNN, SVM, NB and LR base classifiers), three AdaBoost methods (with DT, SVM and NB base classifiers), methods of stochastic design of ensembles of the decision trees (RandomForest and ExtraTrees), gradient boosting and voting classifier.

The accuracy of the proposed TSCM method is at the level of the best techniques (Table 2). Only the ExtraTrees and MLP (40,5) have the comparable characteristics by check results on the test set. It should be noted that the TSCM shows a poorer performance on the training set than some other methods but it exceeds them on the objects recognition accuracy on the test set. It means that it avoids overfitting much more successfully unlike the other methods. Besides this, the TSCM shows rather balanced ratio of precision and recall ratios, i.e. it finds bankrupts (recall=0.953) and not bankrupts (precision=0.910) with rather high exactness. Thus, the TSCM is the best in the studied group of methods in terms of the balance of Type I and Type II errors.

Table 3 shows four examples of how the proposed method can be applied in practice. Consider four companies:

- Gazprom Gaz Distribution Barnaul. This successful company builds and maintains gas distribution networks in the Altai region, is part of the Gazprom group.

Table 3
Examples of method application.

Firm	Classifiers outcomes C_i					C_E	Decision
	KNN	LR	NB	DT	SVM		
Gazprom Gas Distribution Barnaul	–1	–1	–1	–1	–1	–0889	non bankrupt
Vanino Commercial Sea Port	–1	1	–1	–1	1	–0289	non bankrupt
VKZ	1	1	–1	1	1	0645	bankrupt
Ru-Energy Group	1	1	1	1	1	0889	bankrupt

- Vanino Commercial Sea Port. This is the successful stevedore company in the Khabarovsk region, which provides loading and freight forwarding services.
- VKZ. This company produces wall tiles and porcelain stoneware. The Arbitration Court of the Voronezh Region declared it as bankrupt on October 7th, 2014.
- Ru-Energy Group. This is a holding, which included companies that produce oil and gas equipment. The enterprise was declared bankrupt by the Arbitration Court of Moscow on May 15th, 2015 and is currently in the process of liquidation.

According to the proposed method, outcomes C_i of ordinary classifiers trained on the first stage were obtained and their values are presented in Table 3. The final decision of ensemble is calculated as the sum $C_E = \sum_{i \in M} w_i C_i$, where w_i is the weight of i th ordinary classifier defined on the second stage of training. Weights are presented in Table 1, values of C_E are presented in Table 3. As it was defined above, the sign of C_E defines the object class; the value $C_E > 0$ corresponds to bankrupts and the value $C_E < 0$ corresponds to successful companies. The absolute value of C_E corresponds to the degree of confidence.

Table 3 shows that sometimes basic models can misclassify firms (e.g. LR for Vanino Commercial Sea Port and NB for VKZ), but weighted summation suppresses these errors.

It is necessary to note that the proposed method requires more computation time because of GA usage. The time complexity of GA can be estimated as $O(P * G * O(f) * (p_c * O(c) + p_m * O(m)))$, here P is population size; G is number of generations; f , c , and m are functions to calculate fitness, crossover and mutation respectively; p_c and p_m are probabilities of crossover and mutation. Crossover and mutation operations in our case are simple, and their complexity is $O(c) = O(m) = O(1)$. So, overall complexity is $O(P * G * O(f))$, where $O(f)$ is complexity of calculation of accuracy of basic classifier. Therefore, the basic classifiers training time (the first stage) can be estimated as $t_1 = P * G * \sum_{i \in M} t_i$, here t_i is a time to assess the accuracy of i th basic classifier (this time includes fitting on train set, prediction on test set and accuracy evaluation). Similar estimation can be obtained for the second stage of method, but in the last case t_i includes only operations of prediction and accuracy evaluation. So, we can conclude that time required for the first stage of ensemble training according to the proposed method exceeds the training time of the base classifiers approximately by $P * G$ times, in our case ($P=50$ and $G=50$) by 2500 times. Time required for the second stage ($P=40$ and $G=40$) exceeds the accuracy check time for base classifiers by 1600 times. This is not a serious problem, since the training of one classifier with the help of scikit-learn takes little time (less than 0.1s), the speed measurements on a computer with two 1.5 GHz CPU cores have shown that the average training time of the ensemble is about 400 s with consecutive training of each classifier. It is obvious, that this time can be extremely reduced with the parallel training of basic classifiers.

Commonly the main issue of GA is not the time complexity but the convergence rate (Oliveto, He, & Yao, 2007; Rabinovich & Wiggerson, 1999), which is dependent on the number of iterations (i.e. $P * G$). These parameters were chosen experimentally to balance the computational time and the convergence rate.

Table 4

Features of firms in the research data set.

Feature name	Description	kNN	LR	NB	DT	SVM	Total
L1	The coefficient of current liquidity	1	1	1	1	1	5
L2	Quick (urgent) liquidity ratio	1	1	0	1	1	4
L3	Absolute liquidity ratio	0	1	1	1	1	4
L4	The liquidity ratio at raising funds	1	1	0	1	1	4
T1	Receivables turnover ratio	0	1	0	1	1	3
T2	The turnover ratio of funds	1	1	0	1	1	4
T3	Total assets turnover	1	1	0	1	1	4
T4	Net income to net capital ratio	0	1	1	1	1	4
S1	The ratio of own and borrowed funds	0	1	0	1	1	3
S2	Leverage ratio	1	1	1	1	1	5
S3	Retained earnings to total assets	1	1	0	1	1	4
S4	Current assets to total assets	1	1	1	1	1	5
S5	Net working capital to total assets	1	1	0	1	1	4
S6	Ratio of financing	0	1	0	1	1	3
S7	The rate of investment	0	1	0	1	0	2
S8	The ratio of own working capital	1	1	0	1	1	4
S9	The flexibility ratio of own funds	0	1	1	1	1	4
S10	The ratio of EBITDA to interest paid	1	0	1	1	1	4
S11	EBITDA	0	0	1	1	1	3
S12	Net assets	0	1	1	1	1	4
S13	The coefficient of autonomy	1	1	1	1	1	5
S14	The degree of solvency in current liabilities	1	1	0	1	1	4
S15	The ratio of own and borrowed funds	0	1	1	1	1	4
R1	Profitability of sold products	0	1	1	1	1	4
R2	Return on assets (ROA)	0	0	1	0	1	2
R3	EBIT to Total Assets.	0	1	1	1	1	4
R4	Return on net assets by net profit	1	1	0	1	1	4
R5	Profitability	1	1	0	0	0	2
R6	Profitability of sales	1	1	0	1	1	4
I1	Government control	0	1	1	1	1	4
I2	Being under sanctions of foreign States	0	1	0	1	0	2
I3	Unreliable supplier	1	1	1	1	1	5
I4	The ratio of the number of arbitration claims against the firm by the number of claims of the firm to other companies	0	1	1	1	1	4
E1	GDP in % to the previous year	1	1	0	1	1	4
E2	Consumer price index	0	1	0	1	1	3
E3	The product price index of industrial goods	1	1	1	1	1	5
E4	Summary index of product prices (cost services) investment destination	0	1	1	1	1	4
E5	The average monthly accrued wages of organizations employees in % to the previous year	1	1	1	1	1	5
E6	Unemployment rate	0	1	0	1	1	3
E7	The level of overdue accounts payable	0	1	0	1	1	3
E8	Herfindahl-Hirschman Index	0	1	1	0	1	3
E9	The company occupies a significant market share	1	1	1	1	1	5
E10	Natural monopoly	1	1	1	0	1	4
E11	The company occupies a significant market share	1	1	0	0	1	3
E12	The average long-term interest rate of the loan	1	1	0	1	1	4
E13	The growth rate of the US Dollar	0	1	1	0	1	3
E14	The rate of revocation of licenses of credit organizations	1	1	1	1	1	5
E15	MICEX index	1	1	1	0	1	4
E16	The growth of the MICEX index	1	1	0	1	1	4
E17	RTS index	1	1	0	1	1	4
E18	The growth of the RTS index	1	1	1	1	1	5
E19	The price of Brent crude	1	1	0	1	1	4
E20	The increase in the price of Brent crude	1	1	0	1	1	4
E21	HSBC PMI index	1	1	0	1	1	4
E22	Growth of HSBC PMI index	1	1	0	1	1	4

It should be noted that proposed method has yet another important advantage: it allows selecting the optimal features set to solve the problem. The complete list of features which are available in the research data set provided in Table 4. Table 4 also lists information, whether the given feature is significant from the point of view of various classifiers (1 in the corresponding column). This information is obtained from the genome G of classifiers, trained at the first stage. The overall significance of each feature can be judged by summing the data for individual classifiers. It is given in the column “Total”.

As expected, 4 coefficients of company financial condition (current liquidity, financial independence, current assets, equity ratio)

have been included in the most significant 10 factors (for which Total=5 in Table 4). The financial indicators are most important factors predicting the bankruptcy. In fact, the classical models of bankruptcy (Altman, 1968; Beaver, 1966; Fulmer, Moon, Gavin, & Erwin, 1984; Ohlson, 1980; Springate, 1978; Taffler & Tisshaw, 1977; Zmijewski, 1984) use only financial ratios. However, according to our technique, the factors identified, e.g. by Altman (1968), are not included in the 10 most significant. It is consistent with the result of LDA modeling described above. It means that it is impossible to predict the bankruptcy reliably using only financial ratios, especially for developing economies, including Russia. In our case, six environmental factors were also included in the list

Table 5
Testing of feature selection quality.

Features selection rule	Number of features	Training set			Test set		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall
all features	55	0.963	0.943	0.987	0.934	0.910	0.953
Total > 2	51	0.971	0.992	0.989	0.951	0.932	0.965
Total > 3	41	0.956	0.929	0.989	0.929	0.900	0.953
Total > 4	10	0.937	0.909	0.973	0.918	0.880	0.953

of most significant. Among these factors are the changes in the macroeconomic situation: the rate of revocation of licenses of credit institutions, the increase of the RTS index and the price index of industrial goods. These factors act as signs of crisis situations provoking an additional stress for the firms. The importance of the macroeconomic factors for bankrupts forecasting has been proved recently (Alifiah, 2014; Karas & Režňáková, 2014; Tinoco & Wilson, 2013). In this paper, the microeconomic indicators of the external environment were significant as well, although they are not taken into consideration in the other studies so often. Therefore, in order to predict a firm's bankruptcy in Russia, it is crucial to know what part of the market the company occupies, and the fact whether it is on the list of unscrupulous suppliers and level of the average monthly salary.

18 factors of the firm financial condition and 13 factors of the external environment have put together a less important group of 31 factors (Total = 4 in Table 4). Indicators of the stock indexes, oil price and GDP were significant at the macro level. Natural monopoly, the government control and the number of claims were also classified as significant factors. Therefore, we confirm results of numerous studies (e.g., Tinoco & Wilson, 2013) that consolidation of external environment indicators and financial performance provides higher prognostic capability of model. In our research the proposed algorithm selected external environment indicators at the micro level in addition to standard macroeconomic environment indicators (the stock indexes, GDP, inflation, an oil price). The different types of indicators of competition are the level of industry monopolization (whether the firm is a natural monopoly, and the presence of government control), and possible issues with firm suppliers along with the number of claims to the firm. So, it is necessary to consider the wide set of external and internal environment factors for entity bankruptcy forecasting.

Three variants of feature sets were constructed, including 51 (Total > 2 in Table 4), 41 (Total > 3) and 10 (Total > 4) factors to assess the ability of the proposed method to solve the features selection problem. The data of the proposed method testing on different features sets are shown in Table 5.

All characteristics of the method increase in case of exception of features that are significant for less than 50% of the classifiers included in ensemble (the rule of sample is Total > 2). The characteristics degrade in case of further reduction of number of features. Thus, GA based features selection can be used as the preprocessing tool in case of significant number of factors. However, the choice of the rule of insignificant features cutting off shall be researched for each problem.

6. Conclusion and future research

The obtained results allow taking a fresh look at the conclusions of Bellovary et al. (2007) and other authors. Although the methods for bankruptcy prediction based on discriminant analysis and ANN models yield good results, they could be greatly improved by contemporary methods of machine learning. This means that future research should not be confined to already known

models, it is particularly important for studies in developing countries.

The key factor of success is the use of new indicators reflecting the characteristics of the external environment, the specifics of the economic situation, the peculiarities of the management of the firm, the firm size, etc. This increases substantially the predictive ability of models in comparison with use of only financial ratios. However, in this case, the problem of selecting only significant features arises, since larger number of variables does not mean an increase in the predictive ability of model. Our results show that an effective tool for selecting features can be based on genetic algorithms.

In addition, the genetic algorithm is an effective technique for combining different models of ordinary classifiers into the ensemble. In combination with the features selection, it allows to propose a method that exceeds in the accuracy and balance of Type I and Type II errors all other methods.

In our opinion, such approach allows one to design models featuring quick adaptability to the various conditions and bridging the gap between theory and practice.

The disadvantage of this approach is its relatively high computational complexity. This disadvantage, however, is not critical since modern computers have sufficient performance. Tests show that the full retraining of the model takes about 5 minutes even on an ordinary PC.

Regarding the future research direction, there are few important issues to be empirically investigated. First, the proposed method should be tested on data sets of other countries. It is necessary not only to assess its performance in other conditions, the features that will be significant for other countries are of particular interest. This can provide valuable information on the differences in firm behavior and external factors in various countries. Similar studies can be carried out separately for different industries, large and small enterprises, etc.

Another very interesting question is the set of basic classifiers. In this work, we employed only five simple models (kNN, LR, NB, DT, SVM). LDA, QDA and MLP were excluded as they were not yield to the feature selection. More complex classifiers, such as RandomForest and ExtraTrees, can also be used as the base. Our preliminary tests showed that the inclusion of these models in the set of basic classifiers does not improve the quality of forecasting on the Russian data set, but from a practical point of view it is advisable to have as diverse basic set as possible.

The third research direction concerns the comparison of the proposed method with techniques such as blending and stacking, which also support a reduction of feature set (however, in this case the explanatory power is lost) and the combination of base classifiers.

The last but not the least research direction is due to the bankruptcy procedure takes a long time. Therefore, in addition to forecasting the fact of bankruptcy itself, the following research questions arise: (1) forecasting the results of bankruptcy (liquidation, accession, settlements with creditors) and (2) forecasting the duration of bankruptcy procedure.

References

- Acosta-Gonzalez, E., & Fernandez-Rodriguez, F. (2014). Forecasting financial failure of firms via genetic algorithms. *Computational Economics*, 43(2), 133–157.
- Ahn, H., & Kim, K.-J. (2009). Bankruptcy prediction modeling with hybrid case-based reasoning and genetic algorithms approach. *Applied Soft Computing*, 9, 599–607.
- Alifah, M. N. (2014). Prediction of financial distress companies in the trading and services sector in Malaysia using macroeconomic variables. *Procedia - Social and Behavioral Sciences*, 129, 90–98.
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23, 589–609.
- Aziz, M. A., & Dar, H. A. (2006). Predicting corporate bankruptcy: Where we stand? *Corporate Governance: The International Journal of Business in Society*, 6, 18–33.
- Beaver, W. (1966). Financial ratios as predictors of failure. Empirical research in accounting selected studies. *Journal of Accounting Research*, 4, 71–111.
- 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.
- Bredart, X. (2014). Financial distress and corporate governance: The impact of board configuration. *International Business Research*, 7(3), 72–80.
- Brown, G., Wyatt, J., Harris, R., & Yao, X. (2005). Diversity creation methods: A survey and categorisation. *Information Fusion*, 6(1), 5–20.
- Bruneau, C., de Bandt, O., & El Amri, W. (2012). Macroeconomic fluctuations and corporate financial fragility. *Journal of Financial Stability*, 8, 219–235.
- Bryll, R. (2003). Attribute bagging: Improving accuracy of classifier ensembles by using random feature subsets. *Pattern Recognition*, 36(6), 1291–1302.
- Chan, C. Y., Chou, D. W., Lin, J. R., & Liu, F. Y. (2016). The role of corporate governance in forecasting bankruptcy: Pre- and post-SOX enactment. *North American Journal of Economics and Finance*, 35, 166–188.
- Chang, S.-Y., & Yeh, T.-Y. (2012). An artificial immune classifier for credit scoring analysis. *Applied Soft Computing*, 12(2), 611–618.
- Delas, V., Nosova, E., & Yafinovich, O. (2015). Financial security of enterprises. *Procedia Economics and Finance*, 27, 248–266.
- du Jardin, P. (2016). A two-stage classification technique for bankruptcy prediction. *European Journal of Operational Research*, 254, 236–252.
- Duffie, D., Saita, L., & Wang, K. (2007). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics*, 83, 635–665.
- Fedorova, E., Gilenko, E., & Dovzhenko, S. (2013). Bankruptcy prediction for Russian companies: Application of combined classifiers. *Expert Systems with Applications*, 40, 7285–7293.
- Fulmer, J., Moon, J., Gavin, T., & Erwin, M. (1984). A bankruptcy classification model for small firms. *Journal of Commercial Bank Lending*, (July), 25–37.
- Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63, 3–42.
- Gordini, N. (2014). A genetic algorithm approach for SMEs bankruptcy prediction: Empirical evidence from Italy. *Expert Systems with Applications*, 41, 6433–6445.
- Gorzalczany, M. B., & Rudzinski, F. (2015). A multi-objective genetic optimization for fast fuzzy rule-based credit classification with balanced accuracy and interpretability. *Applied Soft Computing*, 40, 206–220.
- Guyon, I., & Elisseeff, A. (2003). An Introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157–1182.
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining: concepts and techniques*. WalthamMA: Morgan Kaufmann.
- Heo, J., & Yang, J. Y. (2014). AdaBoost based bankruptcy forecasting of Korean construction companies. *Applied Soft Computing*, 24, 494–499.
- Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8), 832–844.
- Karas, M., & Režňáková, M. (2014). To what degree is the accuracy of a bankruptcy prediction model affected by the environment? The case of the Baltic States and the Czech Republic. *Procedia - Social and Behavioral Sciences*, 156, 564–568.
- Kim, M.-J., & Kang, D.-K. (2012). Classifiers selection in ensembles using genetic algorithms for bankruptcy prediction. *Expert Systems with Applications*, 39, 9308–9314.
- Kumar, P. R., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques—A review. *European Journal of Operational Research*, 180(1), 1–28.
- Kuncheva, L. (2004). *Combining pattern classifiers: methods and algorithms*. Wiley-Interscience.
- Lee, S., & Choi, W. S. (2013). A multi-industry bankruptcy prediction model using back-propagation neural. *Expert Systems with Applications*, 40, 2941–2946.
- Liang, D., Lu, C.-C., Tsai, C.-F., & Shih, G.-A. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European Journal of Operational Research*, 252, 561–572.
- Liang, D., Tsai, C.-F., & Wu, H.-T. (2015). The effect of feature selection on financial distress prediction. *Knowledge-Based Systems*, 73, 289–297.
- Lee, T.-S., & Yeh, Y.-H. (2004). Corporate governance and financial distress: Evidence from Taiwan. *Corporate Governance: An International Review*, 12(3), 378–388.
- Mukhopadhyay, A., Maulik, U., & Bandyopadhyay, S. C. (2014). A survey of multi-objective evolutionary algorithms for data mining: Part I. *IEEE Transaction on Evolutionary Computation*, 18(1), 4–19.
- Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18, 109–131.
- Oliveto, P. S., He, J., & Yao, X. (2007). Time complexity of evolutionary algorithms for combinatorial optimization: A decade of results. *International Journal of Automation and Computing*, 4(3), 281–293.
- Onan, A., Korukoglu, S., & Bulut, H. (2016). A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification. *Expert Systems With Applications*, 62, 1–16.
- Pedregosa, F. (2011). Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Peng, Y., Wang, G., Kou, G., & Shi, Y. (2011). An empirical study of classification algorithm evaluation for financial risk prediction. *Applied Soft Computing*, 11(2), 2906–2915.
- Rabinovich, Y., & Wigderson, A. (1999). Techniques for bounding the convergence rate of genetic algorithms. *Random Structures Algorithms*, 14(2), 111–138.
- Rowoldt, M., & Starke, D. (2016). The role of governments in hostile takeovers – Evidence from regulation, anti – takeover provisions and government interventions. *International review of Law and Economics*, 47, 1–15.
- Salloum, C., & Azoury, N. (2012). Corporate governance and firms in financial distress: Evidence from a Middle Eastern country. *International Journal of Business Governance and Ethics*, 7(1), 1–17.
- Santos, Dos, E. M., Sabourin, & Maupin, P. (2009). Overfitting cautious selection of classifier ensembles with genetic algorithms. *Information Fusion*, 10, 150–162.
- Sartori, F., Mazzucchelli, A., & Di Gregorio, A. (2016). Bankruptcy forecasting using case-based reasoning: The CREPERIE approach. *Expert Systems With Applications*, 64, 400–411.
- Springate, G. L. (1978). *Predicting the possibility of failure in a canadian firm*. Simon Fraser University.
- Taffler, R. J., & Tisshaw, H. J. (1977). Going, going, gone, four factors which predict. *Accountancy*, 88, 50–54.
- Tinoco, M. H., & Wilson, N. (2013). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *International Review of Financial Analysis*, 30, 394–419.
- Thammasiri, D., & Meesad, P. (2012). Ensemble data classification based on diversity of classifiers optimized by genetic algorithm. *Advanced Materials research*, 433, 6572–6578.
- Tsai, C. F. (2014). Combining cluster analysis with classifier ensembles to predict financial distress. *Information Fusion*, 16, 46–58.
- Tsai, C.-F., Hsu, Y.-F., & Yen, D. C. (2014). A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing*, 24, 977–984.
- Vlami, P. (2007). Default risk of UK real estate companies: Is there a macro-economy effect? *The Journal of Economic Asymmetries*, 4(2), 99–117.
- Wozniak, M., Grana, M., & Corchado, E. (2014). A survey of multiple classifier systems as hybrid systems. *Information Fusion*, 16(1), 3–17.
- Zhou, L., Lai, K. K., & Yen, J. (2014). Bankruptcy prediction using SVM models with a new approach to combine features selection and parameter optimization. *International Journal of System Science*, 45(3), 241–253.
- Zięba, M., Tomczak, S. K., & Tomczak, J. M. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Systems With Applications*, 58, 93–101.
- Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22, 59–82.