Sinan Aktan (Germany)

Application of machine learning algorithms for business failure prediction

Abstract

Business failure prediction has long been an active research field in finance. Due to advent of new regulations like Basel II business failure prediction methods moves from traditional statistical models to more comprehensive machine learning techniques. In this context the article investigates efficiency of 8 machine learning algorithms such as Naive Bayes, Bayesian Network, k-NN, ANN, SVM, C4.5, CHAID and CRT in financial distress. For cost sensitive prediction variables selected through two variable elimination phases of ANOVA and cost sensitive attribute evaluator algorithm embedded in WEKA platform. For performance evaluation not only classification accuracy but also AUROC values are taken into consideration. CRT outperforms all other learning algorithms; CHAID fails to produce significant classification for three annual periods prior to failure occurrence time. Except CRT all other learning algorithms are superior to each other in terms of classification accuracy and AUROC.

Keywords: financial distress prediction, bankruptcy prediction, machine learning, cost sensitive, financial ratios.
JEL Classification: C11, C13, C14, C45, G32, G33.

Introduction

Business failure prediction has long been important and studied widely by financial literature under the name of bankruptcy prediction, firm failure prediction and financial distress prediction. This subject involves developing models that attempt to forecast financial failure before it actually happens, because a failed firm could trigger a contagious loss to financial system involving social and economic losses in national and global dimension. Therefore, the recent default and bankruptcies of many companies and ongoing adaptation process of Basel II, briefly an international standard to adjust the amount of reserve capital to be put aside to prevent banks against types of risks, has further underlined the importance of failure prediction in academia, industry and credit institutions. So measuring credit risk accurately allows credit institutions to achieve targeted return and risk characteristics.

The pioneering study in the field of bankruptcy prediction was conducted by Beaver in 1966. Beaver made the first study in bankruptcies and estimating failure risk of companies. The only point where Beaver was mostly criticized was that his study was dependent on univariate analysis and considered certain groups (a limited number) of financial ratios. In 1968, Altman expanded this analysis to multivariate discriminant analysis.

In bankruptcy prediction studies two main approaches can be distinguished. The first and the most often used approach is the empirical search for predictors that lead to the lowest misclassification rates. The second approach concentrates on the search for statistical and structural methods that would also lead to improved prediction accuracy (Back et al., 1996).

Statistical models have certain distributional hypothesis that financial statement data do not always fit. Hence some non-parametric techniques have been developed to overcome the constraints of traditional statistical models. Most of them belong to data mining domain such as artificial intelligence. Most of the researchers dealt with the issue of comparing data mining methods with traditional statistical models.

In this context, this study could be included partly in this research line. This research will present eight machine learning algorithms: Bayesian Network and Naive Bayes from Bayesian algorithms, k-Nearest Neighbor (k-NN) instance based learning, Artificial Neural Network (ANN) with Multilayer Perceptron (MLP), Support Vector Machine (SVM), C4.5, CHAID and CRT from decision tree algorithms for financial distress classification modeling.

The remainder of research is organized as follows. Section 1 dedicated to brief explanations of the applied models. Prior researches on financial distress classification based on data mining methods are briefly reviewed in Section 2. Section 3 is reserved for data set and variable selection, empirical study and outcomes discussed in Section 4, and the final Section contains the summary and conclusion.

1. Employed data mining techniques

1.1. Bayesian models. The Naive Bayes classifier method is based on the so-called Bayesian theorem, naive assumes independence. The Naive Bayes classifier produces probability estimates rather than predictions. The probability estimate is the conditional probability distribution of the values of the class attribute based on the values of other attributes. In this way Naive Bayes classifier is just an alternative way of representing a conditional probability distribution and can only represent simple distributions (Witten and Frank, 2005). But Bayesian Network is a theoretically well-founded way of representing probability distributions concisely and comprehensibly in a graphical manner. The structure of Bayesian Network (BN) is represented by a directed acyclic graph (DAG), which is a network of nodes; represents attributes, connected by directed edges, expresses dependencies between attributes, in such a way that there are no cycles.

BNs don't suffer from the underlying distributions of variables. BNs don't suffer from missing attributes of instances; instances that have missing variables can be used to train or test BNs. In fact bankrupt and financially distressed firms tend to have missing variables for bankruptcy studies. BNs are dynamic and interactive. BNs can be updated with new information added to the training set and BNs are more transparent and intuitive compared to neural networks because the relationship among attributes are explicitly represented by the DAG (Sun and Shenoy, 2007).

1.2. K-Nearest Neighbor. K-Nearest Neighbor (k-NN) algorithm is one of the most fundamental and simple classification method based on closest training examples in the feature space. K-NN is a type of instance based algorithm in the category of lazy learning algorithm (Aha, 1997). K-NN classifies an object based on its similarity to other objects. The logic assumes similar objects are near each other and dissimilar objects are distant from each other. So an object is labeled according to label of majority of its neighbors. The similarity of objects is assessed by using suitable distance metric, usually Euclidian distance is used as a distance metric for continuous variables. However, there is not a common concept of defining number of nearest neighbor, researcher sets it in order to have good classification accuracy; but it makes intuitive sense to use more than one nearest neighbor if the size of training set is large.

This simple method has some practical problems, it tends to be slow for large training set, it performs badly with noisy data and it performs badly with irrelevant attributes because each attribute has the same influence on the decision, just as it does in the Naive Bayes method (Witten and Frank, 2005). On the other hand the advantage of this simple method over most other machine learning methods is that it allows adding new examples to training set at anytime.

1.3. Artificial Neural Networks. Artificial Neural Networks (ANN) is another machine learning tool based on computational models inspired from biological network of neurons found in human central nervous system. The most prominent ANN algorithm in the financial distress prediction domain is Multi Layer Perceptron (MLP), which is composed of three layers; input layer contains the predictors namely attributes, hidden layer contains the unobservable nodes, the output layer contains the responses, and there can be several hidden layers for complex applications. The most frequently used algorithm for learning MLP is the Back Propagation algorithm (BPA). BPA uses gradient descent which can find local minimum, if the function has several minima, for MLP has many, it may not find the best one. This is a significant drawback for standard MLP compared with Support Sector Machine (SVM) (Witten and Frank, 2005).

ANN is more adaptive to real world situation it could discriminate non-linear patterns, so it does not suffer from constraints of statistical models. However ANN has several drawbacks, it is a black box procedure, and it is hard to interpret the results due to lack of explanatory power and lack of feature selection, it needs too much time and efforts to construct a best architecture (Lee, 2006).

1.4. Support Vector Machines. Support Vector Machine (SVM) was introduced by Vapnik (1995). It is a blend of linear modeling and instance based learning, it selects a small number of critical boundary instances called support vectors from each class and builds a linear discriminant function that separates each class as wide as possible. The system transcends the limitations of linear boundaries by making it practical to include nonlinear terms in the function, making it possible to form quadratic, cubic and higher order decision boundaries.

The basic idea of SVM is to use linear model to implement nonlinear class boundaries through some nonlinear mapping the input vector into the high dimensional feature space. A 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 is constructed. Thus SVM is known as the algorithm that finds a special kind of linear model, the maximum margin hyperplane. The maximum margin hyperplane gives the maximum separation between the decision classes. The training examples that are closest to maximum margin hyperplane are called support vectors. All other training examples are irrelevant for defining binary call boundaries.

Support vector machines like neural networks are not suffering from constraints of statistical distributions. With support vector machines it is unlikely to occur overfitting and they often produce very accurate classifiers. On the contrary, the computation is very complex and they are slow compared to other machine learning algorithms when applied in nonlinear setting.

1.5. Decision trees. Decision tree is the implementation of divide and conquer strategy to set of independent instances to learn the problem. Decision tree is composed of root, internal decision nodes and terminal leaves. Each node in a decision node represents a test of a particular attribute or a function of one or more attributes in the instance set to be classified. The outcome of test represents branches so each branch represents the test value that the node can take. This process starts at root and is repeated recursively until a leaf node is reached then the instance is classified according to class assigned to the leaf.

There are different types of decision tree algorithms which are ID3, CRT (Classification and Regression Tree) and CHAID (Chi-Square Automatic Interaction Detector). ID3 was introduced by J. Ross Quinlan in 1979, ID3 was later enhanced in the version C4.5 and now C5.0. ID3 and enhanced algorithms split the attributes based on the gain in information that the split provides. CRT and CHAID are relatively new and popular non-parametric analysis techniques. CRT algorithm builds decision tree using gini, towing or ordered towing criterion to choose the optimum split, whereas CHAID algorithm uses chi-square statistics for optimum splits.

Decision tree is a nonlinear architecture able to discriminate nonlinear patterns and doesn't suffer from any distributions constraints. It is easy to interpret the results, doesn't require too much time preparation of initial data and performs well large data.

2. Prior research

Financial distress prediction is of particular interest to accounting and finance for a long time. Systematic studies in this domain gained density after Beaver's (1966) univariate and Altman's (1968) multivariate analysis. Most of the later studies tried to develop further Altman's model or establish alternative models. Until the 1980s DA (Discriminant Analysis) was the dominant method in failure prediction. The first practitioner of logit analysis in the failure prediction was Ohlson (1980). Most of the studies conducted after 1981 used logit analysis to relax the constraints of DA.

In fact, the constraints of traditional statistics were always a discussion point and criticized heavily, so this circumstance motivated the practitioners to switch into structural financial forecasting models (which are not in the scope of this study) and nonparametric models. Some of the nonparametric studies could be summarized as follows.

Marais et al. (1984) applied RPA (recursive partitioning algorithm) for modeling commercial bank loan classification and compared the model with probit. They found that RPA was not significantly better than probit.

Frydman et al. (1985) used RPA and DA in financial distress prediction. Less complex RPA model was found to perform better than DA in terms of cross-validated and bootstrapped accuracies.

Messier and Hansen (1988) used inductive algorithm ID3 in loan default and bankruptcy prediction. The results were evaluated by comparing with the results of DA. ID3 outperformed DA on the other hand both models had partly common predictive attributes.

Odom and Sharda (1990) developed neural network model for bankruptcy prediction and compared the results with that of DA in terms of classification accuracy. They asserted that neural networks might be used in bankruptcy prediction domain.

Cronan et al. (1991) applied RPA to datasets representing the mortgage, commercial, and consumer lending problems and compared the results with that of DA, logit, probit, and ID3. RPA provided superior results than ID3 and other statistical models while using fewer variables.

Tam and Kiang (1992) applied data mining in bank failure prediction and they used ANN as main model and compared with DA, logistic regression, k-NN and ID3 in terms of prediction accuracy, adaptability, and robustness. Back propagation network outperformed other models. Statistical models found better than ID3 and k-NN which was the least accurate model.

Coats and Font (1993) built neural network to estimate the future financial health of firms. Neural network used for identifying data patterns that distinguish healthy firms from distressed ones. Their results suggested that neural network approach was more effective than DA.

Godwin Udo (1993) built neural network model to predict going concern of firms based on financial ratios, the results indicated that neural network was more accurate than multiple regression analysis.

Wilson and Sharda (1994) compared the prediction capabilities of neural network model and DA model. They found out that the result of NN model was significantly superior to DA in bankruptcy prediction.

Altman et al. (1994) applied neural network on Italian Centrale dei Bilanci's dataset consists of over 1000 Italian firms and compared the results with that of DA. The results indicated that both model provided balanced classification accuracy. They suggested that both model could be combined for predictive reinforcement.

Boritz and Kennedy (1995) examined two neural network approaches, Back-Propagation and Optimal Estimation Theory, for predicting bankruptcy filing. The model based on Optimal Estimation Theory approach had the lowest Type I error and highest Type II error while traditional statistical techniques DA, logit and probit had the reverse relationship. The model based on Back-Propagation approach had the intermediate level of Type I and Type II errors. The results indicated that performance of the models were sensitive to the selected variables.

Back et al. (1996) applied DA, logit and genetic algorithms to find out predictors of bankruptcy. The result revealed that all of the models chose different number of variables as predictors. Logit analysis chose the subset of variables of DA with one variable exception. Neural network chose relatively far more variables than of logit and DA whereas neural network was superior to both statistical models in terms of classification accuracy for one to three years prior to bankruptcy periods.

Henly and Hand (1996) used k-NN method with an adjusted Euclidian distance metric in assessing credit scoring problem. It was found that k-NN performed well in achieving the lowest expected bad risk rate compared to linear regression, logit, decision trees and decision graphs. They asserted that k-NN was prosperous tool for assessing credit score.

Etherige and Sriram (1997) used two ANN models, categorical learning NN and probabilistic NN, and compared them with statistical DA and logit models to examine financial distress one to tree years prior to failure in comparing overall classification error DA and logit outperformed NN models. In fact when relative error cost was considered ANN models performed better than statistical models. The results indicated that ANN models performance increases as the time period moves farther away from the eventual failure date.

Joos et al. (1998) compared the performances of decision tree and logit analysis in a credit classification environment. For this purpose they used extensive database of one of the largest Belgium bank. They asserted that logit models were consistent in credit decision process, on the other hand for the qualitative and short scheme data decision tree models were better.

Varetto (1998) analyzed the comparison of genetic algorithm GA and linear DA. The analysis was conducted to 1920 sound and counterparty mate companies to assess insolvency risk. He concluded that the analysis proved that GA was effective method for insolvency diagnosis although the result of LDA was superior to GA.

Yang et al. (1999) applied probabilistic NN instead of Back-Propagation NN for bankruptcy prediction and compared the results with that of DA. They asserted that probabilistic NN without pattern normalization and Fisher DA provided best overall estimation, but DA produced outstanding results for bankrupt companies.

Lin and McClean (2000) used four classification models DA, logit, NN, and DT for prediction of financial distress. Each model was subject to three variable selection methods, human judgment, ANOVA and factor analysis. They found that the variables selected by ANOVA provided better results and among classifier models DT and NN outperformed statistical models in terms of classification accuracy.

Ko et al. (2001) used Liang's CRIS (composite rule induction system) model and compared with NN and logit model in corporate financial distress prediction domain. They asserted that CRIS and NN outperformed logit model; however, despite the higher performance of CRIS and NN, the extracted rules by CRIS are easier to be understand by the human.

Atiya (2001) was inspired by Merton's asset value model so brought new variables, extracted from stock price, in domain of bankruptcy prediction. He showed that using market based variables in addition to traditional financial ratio variables resulted in significant increase of classification accuracy by 4 % for three year prior to bankruptcy.

Sarkar and Siriram (2001) developed Bayesian network (BN) models to help human auditors in assessing bank failures. Their Naive Bayesian network and composite attribute BN's performance in classification accuracy was comparable to DT algorithm C4.5. They underlined that the sharpness of BN increases when recent financial indicators are used in models.

Park and Han (2002) introduced an Analytic Hierarchy Process weighted k-NN model, a derivative of k-NN method in bankruptcy prediction area, and compared the performance of new model with regression, logit, weighted k-NN and pure k-NN. The results were in favor of AHP weighted k-NN.

Yip (2003) introduced a hybrid CBR model that uses statistical evaluation for automatically assigning attribute weights and nearest neighbor algorithm for case retrieval. Comparison with DA proved that the model would be competitive alternative in failure prediction context while it outperformed traditional statistical model.

Härdle et al. (2004) implemented SVM for corporate bankruptcy prediction and compared it with DA. SVM outperformed DA slightly in terms of classification accuracy; however the difference was not significant at 5%. On the other hand they proved that SVM was capable to extract information from real life economic data sets.

Shin et al. (2005) used SVM with RBF (Radial Basis Function) for bankruptcy prediction on mid-sized Korean manufacturing firms' dataset. They asserted that lower value of upper bound parameter C leads model to underfit data, per contra large values of C indicates overfit; whereas lower values of kernel parameter δ leads to overfit data, on the contrary higher values indicates the inclination to underfit. Their best values for (C, δ) were (75, 25) and the classification accuracy was superior to BPN. They concluded that there was no systematic way to define optimum kernel function parameters.

Min and Lee (2005) applied SVM for bankruptcy prediction by utilizing 5-fold cross-validation and grid search for optimal parameters of upper bound C and kernel parameter δ for RBF. The found optimal values with cross-validation for (C, δ) were (2^{11}, 2^{-7}). They tested model's classification accuracy by comparing with BPN, DA, and logit. The SVM model found superior against other models. They underlined that there was no common way to define the values of the parameters and which kernel function to use.

Kotsiantis et al. (2005) investigated efficiency of machine learning techniques in the domain of bankruptcy prediction. In this regard Naive Bayes, C4.5, Local Decision Stump, Ripper and RBF algorithms were trained using 150 failed and solvent Greek firms. The result indicated that machine learning algorithms could enable analyst to predict bankruptcy with satisfactory accuracy long before bankruptcy.

Hu and Ansell (2006) studied financial distress prediction with five credit scoring techniques, NB, logit, RPA, ANN and SVM with SMO (sequential minimal optimization). They conducted the study considering the USA, European and Japanese retail market. All market models presented best classification accuracy for one year prior to financial distress. The US market model performed relatively better than European and Japanese models for five years prior to financial distress. In regard to constructed composite model compared to Moody's credit ratings, SVM was the best performing model closely followed by ANN, logit model was the least performing model similar to Moody's.

Lee (2006) introduced Genetic Programming DT model which is integration of GP and DT with C4.5 where functions used in GP are attributes of DT. This integration facilitated DT builder model to handle incremental training data, in other words GP could be considered as DT breeder. GD-DT found to be superior to CART, C5.0, ANN and logit in terms of classification accuracy and AUROC (area under the ROC curve).

Kirkos et al. (2007) explored the effectiveness of data mining classification techniques in detecting firms issuing fraudulent financial statements. In connection with detecting fraudulent financial statements, DT, NN and Bayesian Belief Networks were employed. Bayesian Belief Network showed best performance in terms of classification.

Zheng and Yanhui (2007) used CHAID algorithm for corporate financial distress prediction and compared the results with that of ANN model. The results indicated that CHAID decision tree model is capable to predict financial distress with providing interpretable classification figures.

Auria and Moro (2008) used SVM for solvency analysis and compared the prediction accuracy with that of logistic regression and DA. They mentioned that the performance of SVM model improved by integration of nonlinear separable variable to four financial variable based SVM. Those four variables used for company rating by Deutsche Bundesbank. They used company data provided by Deutsche Bundesbank. Their best model revealed with (C, δ) as (10, 4) and (10, 2,5) for manufacturing and trade sector respectively. Their analysis also showed the lack of systematic method to define kernel function parameters.

Quintana et al. (2008) used Evolutionary Nearest Neighbor Prototype Classifier (ENPC), which is an evolutionary nearest neighbor algorithm, in bankruptcy prediction domain and it received good results compared to other machine learning algorithms NB, logit, C4.5, PART (builds partial C4.5), SVM, and ANN with MLP in terms of classification accuracy. They asserted that ENPC algorithm could be considered an alternative method for bankruptcy prediction.

Lin et al. (2009) constructed a hybrid model using Rough Set Theory (RST), Grey Rational Analysis (GRA) and CBR for business failure prediction. They used RST as preprocessing for relevant attribute selection, then used GRA to derive attribute weights for CBR retrieval process. This hybrid model produced better classification accuracy than RST-CBR (with equal weights) and CBR itself.

Vieria et al. (2009) analyzed financial distress with SVM, NN with MLP (multi layer perceptron) and Addaboost M1 using DIANE database of small and medium size French companies. Constructed models compared with logit analysis in terms of prediction accuracy. SVM achieved the highest accuracy but all models showed comparable results. They stressed that large sets of inputs in classifier can reduce both error types.

Aghaie and Saeedi (2009) aimed to construct financial distress prediction model based on Bayesian Networks. They tested the model with the variables revealed by two different variable choosing methods, conditional correlation between variables and conditional likelihood respectively. The model with variables chosen through conditional likelihood performed slightly better, on the other hand the other BN produced the same classification accuracy as logistic regression did. They claimed that BN could be used alternative method for financial distress prediction. Moreover, they found that companies having lower profitability, more long-term liabilities and lower liquidity are more inclined to financial distress.

Derelioglu et al. (2009) used NN with MLP for SME's credit risk analysis. The conducted model was compared with k-NN and SVM. The variables of the models have been chosen DT, Recursive Feature Extraction (RFE), factor analysis, and principal component analysis. The NN model produced slightly better results than other models.

Koyuncugil and Ogulbas (2009) aimed to construct a data mining model for detecting financial and operating risk indicators of financial distress, the chosen algorithm for modeling purpose was CHAID which is supposed to be easy to understand, easy to interpret and easy to apply by non-professionals of SME. Financial ratios derived from financial tables and the operational variables extracted by questionnaire distributed to SME's located in OSTIM Organized Industrial Zone in Ankara. The study has not been completed yet. After completion of the study, the constructed model will be turned into software for SME's.

Empirical study is carried out by Microsoft Excel, SPSS 15 for windows and WEKA 3.6 open source machine learning software developed at WEKA, The University of Waikato.

3. Empirical design

3.1. Sample selection. Initial sample is composed of 180 production industry firms quoted to ISE with 150 non-distressed and 30 financially distressed firms in 2001 just after the crisis period.

Financially distress firms are defined by two criterions:

1. Turkish Bankruptcy Law article 179 pursuant to Turkish Trade Law articles 324 and 434; business enterprises incurring 2/3 loss in capital stock could be defined as bankrupt. Bankruptcy is a legal procedure, even though those companies selected according to this criteria did not bankrupt, those companies could be classified as financially distressed.
2. Negative equity figures.

In this study, for the initial sample, the ratios are derived from financial statements dated one annual reporting period prior to financial distress occurrence. The data (financial statements) were derived from Istanbul Stock Exchange (www.imkb.gov.tr).

3.2. Variable selection. After the initial groups are defined and firms selected, balance sheet and income statement data are collected. 53 financial ratios have been found useful for this study. 26 financial ratios of variable set have been used in discriminant models of Beaver's (1966) univariate analysis and multivariate analysis of Altman (1968), Deakin (1972), Edminster (1972), Blum (1974), Altman et al. (1977), and El Hennawy and Moris (1983) which are representative examples of studies that used multiple discriminant analysis technique. Moreover, additional 27 financial ratios from independent investment investigation company IBS analysis (www. analiz.ibsyazilim.com) have been found useful for this study. These variables are classified into 6 standard ratio categories. In Table 1 aggregate financial ratios, their codes and ratio categories are presented.

Table 1. Aggregate financial ratios found to be useful

Ratio category	Ratios	Ratio code	Analysts
Liquidity ratios	Current ratio	Lq1	B, D, A-H-N
Liquidity ratios	Quick ratio	Lq2	D
Liquidity ratios	Cash ratio	Lq3	E, D
Liquidity ratios	Working capital to Total assets ratio	Lq4	B, A, D
Liquidity ratios	Current assets to Total assets ratio	Lq5	D, E-M
Liquidity ratios	Quick Assets to Total assets ratio	Lq6	D, E-M
Liquidity ratios	Quick Assets to Inventory ratio	Lq7	B*
Liquidity ratios	Cash to Total assets ratio	Lq8	D

Table 1 (cont.). Aggregate financial ratios found to be useful

Ratio category	Ratios	Ratio code	Analysts
Liquidity ratios	Cash flow to Short-term debts ratio	Lq9	E
Liquidity ratios	Cash flow to Total assets ratio	Lq10	E-M
Liquidity ratios	Cash flow to Total debts ratio	Lq11	B*, B, D
Liquidity ratios	Working capital to Equity ratio	Lq12	IBS
Leverage ratios	Total debts to Total assets ratio	Lv1	B, D
Leverage ratios	Short-term debts to Total assets ratio	Lv2	IBS
Leverage ratios	Short-term debts to Total debts ratio	Lv3	IBS
Leverage ratios	Long-term debts to Total assets ratio	Lv4	IBS
Leverage ratios	Financial debts to Total assets ratio	Lv5	IBS
Leverage ratios	Interest coverage ratio	Lv6	A-H-N
Leverage ratios	Long-term debts to Equity ratio	Lv7	E-M
leverage ratios	Short-term debts to Equity ratio	Lv8	E
Leverage ratios	Total debts to Equity ratio	Lv9	IBS
Fiscal structure ratios	Tangible fixed assets to Long-term debts ratio	Fs1	IBS
Fiscal structure ratios	Equity to Fixed assets ratio	Fs2	IBS
Fiscal structure ratios	Fixed assets to Long-term debts ratio	Fs3	IBS
Fiscal structure ratios	Financial fixed assets to Fixed assets ratio	Fs4	IBS
Fiscal structure ratios	Financial fixed assets to Long-term debts ratio	Fs5	IBS
Fiscal structure ratios	Retained earnings to Total assets ratio	Fs6	A, A-H-N
Activity ratios	Account receivable turnover ratio	A1	IBS
Activity ratios	Inventory to Net sales ratio	A2	E
Activity ratios	Payables turnover ratio	A3	IBS
Activity ratios	Net working capital to Net sales ratio	A4	E, D
Activity ratios	Current assets to Net sales ratio	A5	D
Activity ratios	Tangible fixed assets turnover ratio	A6	IBS
Activity ratios	Total assets turnover ratio	A7	A
Activity ratios	Long-term debt turnover ratio	A8	IBS
Activity ratios	Equity to Net sales ratio	A9	E
Activity ratios	Quick assets to Net sales ratio	A10	D
Activity ratios	Cash to Net sales ratio	A11	D
Profitability ratios	Gross profit margin	P1	IBS
Profitability ratios	Net profit margin	P2	IBS
Profitability ratios	Operational profit margin	P3	IBS
Profitability ratios	Operating profit margin	P4	IBS
Profitability ratios	EBIT margin	P5	IBS
Profitability ratios	Taxes to Net sales ratio	P6	IBS
Profitability ratios	Taxes to Profit before taxes ratio	P7	IBS
Profitability ratios	Return on equity	P8	IBS
Profitability ratios	Return on long-term debts	P9	IBS
Profitability ratios	Return on assets	P10	B, D
Profitability ratios	Financial expenses to Inventories ratio	P11	IBS
Profitability ratios	EBIT to Total assets ratio	P12	IBS
Profitability ratios	Operating income to Total assets ratio	P13	A, A-H-N
Market value ratio	Market to book ratio	M1	IBS
Market value ratio	MV of equity to Book value of debts ratio	M2	A, A-H-N

Notes: A (Altman, 1968), A-H-N (Altman, Haldeman and Narayanan, 1977), B (Beaver, 1966), B* (Blum, 1974), D (Deakin, 1972), E (Edminster, 1972), E-M (El Hennawy and Morris, 1983), IBS (IBS analysis).

The sample selection method of this study follows the same pattern of financial failure studies in international literature. Those studies consider 3 or 5 annual periods prior to failure occurrence of each firm. Each annual period prior to failure occurrence can be represented as -1, -2, -3 and so on; for example, -1 is one annual period prior to failure; -2 is two annual period prior to failure. In this study applied 8 data mining methods' classification power tested at each 3 period.

Variables, the financial ratios that are to be used in the analysis, are selected through two variable elimination stages. In the first stage one-way ANOVA test is conducted. The aim is to define financial ratios of distressed and non-distressed groups that differentiate at 5% significance level. In the second stage the remained variables are put into attribute selection algorithm, which is embedded in WEKA platform, for further elimination.

The outcome of stage 1, the ANOVA test statistics; mean, standard deviation, F-test and its significance level for distressed and non-distressed firms are presented in Table 2. Small significance level indicate group mean differences, in our case the selected 35 financial ratios have significance level less than 5% that means one of the group differs from the other group. The ratios are sorted according to their significance level.

Table 2. ANOVA test statistics

Ratios	Non-distressed		Distressed		Test statistics	
	Mean	Std. dev.	Mean	Std. dev.	F	Sig.
Lv1	0,571	0,203	1,614	1,245	94,560	0,000
P10	-0,012	0,092	-0,578	0,689	93,894	0,000
P13	0,004	0,111	-0,539	0,696	82,706	0,000
Fs2	1,410	1,341	-1,090	2,096	79,951	0,000
Lv5	0,271	0,206	1,075	1,101	69,781	0,000
Lq4	0,170	0,181	-0,701	1,238	68,102	0,000
Lv2	0,441	0,185	1,217	1,112	65,519	0,000
Lq1	1,657	0,927	0,641	0,443	41,890	0,000
Lv4	0,131	0,112	0,397	0,476	38,156	0,000
Lq2	1,099	0,738	0,401	0,352	31,250	0,000
P12	0,135	0,104	0,002	0,239	25,828	0,000
Lq10	0,082	0,101	-0,039	0,240	21,685	0,000
Lq11	0,170	0,214	-0,003	0,153	21,475	0,000
Lq9	0,220	0,269	0,005	0,190	21,083	0,000
P9	0,200	5,402	-5,898	16,480	14,248	0,000
P5	0,288	0,350	-0,996	4,210	13,535	0,000
M2	2,305	2,550	0,717	1,330	13,386	0,000
Lq8	0,096	0,111	0,029	0,047	13,227	0,000
P3	0,112	0,257	-0,858	3,336	12,330	0,001
Lq3	0,341	0,571	0,042	0,088	10,034	0,002
A9	0,376	0,468	2,420	8,109	9,371	0,003
A4	0,379	0,983	-66,019	278,229	8,510	0,004
P8	-0,154	0,422	0,793	3,959	8,167	0,005
Lq6	0,400	0,160	0,312	0,204	7,948	0,005
P2	-0,029	0,277	-27,368	122,974	7,386	0,007
Lq5	0,611	0,169	0,516	0,261	7,305	0,008
P4	0,011	0,355	-27,024	122,641	7,262	0,008
A3	6,494	7,867	2,950	3,228	7,181	0,008
P6	0,034	0,080	0,000	0,000	6,499	0,012
Lv7	0,510	0,761	-0,819	6,193	6,485	0,012
A5	1,329	1,269	3,224	9,194	5,891	0,016
P7	0,230	0,584	0,000	0,000	5,712	0,018
A2	0,363	0,292	1,281	4,757	5,485	0,020
P11	1,156	2,510	9947,342	60477,586	4,042	0,046
Lq7	4,362	10,866	199,541	1196,698	3,974	0,048
A8	14,699	35,687	3,473	4,838	3,630	0,058
Fs6	0,074	0,066	0,049	0,097	3,492	0,063
A7	0,595	0,373	0,469	0,434	3,137	0,078
Lv9	2,184	2,289	-0,919	22,403	2,751	0,099
A1	2,613	1,637	3,210	4,134	1,904	0,169
P1	0,290	0,163	0,236	0,383	1,724	0,191
Fs4	0,106	0,164	0,148	0,239	1,571	0,212
Lv8	1,674	1,784	-0,100	16,954	1,566	0,212
A10	0,895	1,172	1,223	2,254	1,517	0,220
M1	0,961	0,804	0,786	1,314	1,045	0,308
Fs1	6,174	10,480	4,170	12,420	1,001	0,318
Fs3	7,380	12,060	5,285	14,961	0,806	0,370
Lv3	0,773	0,159	0,746	0,208	0,722	0,397

Table 2 (cont.). ANOVA test statistics

Ratios	Non-distressed		Distressed		Test statistics	
	Mean	Std. dev.	Mean	Std. dev.	F	Sig.
Lq12	3,739	3,640	1,493	32,029	0,698	0,404
A11	0,264	0,894	0,157	0,604	0,479	0,490
Lv6	401,854	4248,338	-5,653	35,442	0,339	0,561
A6	4,135	13,480	4,930	12,253	0,107	0,745
Fs5	0,956	2,577	0,821	2,457	0,083	0,774

In this study it is assumed that misclassification errors are not equally important. Cost of Type I error is higher than the cost of Type II error for a credit institution. For example, the model could classify a financially distressed company as non-distressed. This is referred to Type I error, the cost of this error to credit institute would be loss of interest and principle in case of default, and probable recovery costs in a bankruptcy proceedings. On the other hand, the model could classify a non-distressed company as distressed. This is referred to as Type II error, the cost of this error to credit institution would be loss of profit. As a matter, of course, accurate estimation of distressed firms becomes important.

For the cost sensitive modeling purpose in the second stage of variable elimination phase cost sensitive attribute evaluator algorithm, which is embedded in WEKA platform, employed. The reliefF (recursive elimination of features) attribute evaluator is the selected base evaluator of cost sensitive evaluator algorithm, which evaluates the worth of an attribute by repeatedly sampling an instance and considering the value of the given attribute for the nearest instance of the same and different classes. This evaluator can operate on both discrete and continuous class data. The used cost matrix, which is an essential parameter in cost sensitive attribute evaluation, depicted in Table 3 where algorithm weights misclassification of distressed company 10-fold more than misclassification of non-distressed company.

Table 3. Cost matrix

	Non-distressed	Distressed
Non-distressed	0	1
Distressed	10	0

Cost sensitive attribute evaluator ranks the attributes according to their individual evaluations, the selected best 10 variable used for classification modeling listed below:

1. Lq8 – Cash to Total assets ratio;
2. Lq6 – Quick assets to Total assets ratio;
3. Lv5 – Financial debts to Total assets ratio;
4. A2 – Inventory to Net sales ratio;
5. Lq5 – Current assets to Total assets ratio;
6. Lv1 – Total debts to Total assets ratio;
7. Lv2 – Short-term debts to Total assets ratio;
8. P10 – Return on assets,
9. P13 – Operating income to Total assets ratio,
10. Lq10 – Cash flow to Total assets ratio.

Majority of the selected variables belong to liquidity and leverage ratio groups.

Given the limited available sample size for the study it is preferred to employ all the data for training and validation. Nevermore, to avoid probable over-fitting problem, 10-fold cross-validation process is applied. Eventually there is no unique way to define number of folds to be formed, however 10-fold cross-validation preferred commonly by the practitioners.

4. Empirical design and evaluation of models

Empirical study is designed to present and discuss the outcomes of 8 data mining classification models under 5 headings. These classification models are Naive Bayes and Bayesian Network represents the Bayesian models family, k-NN, ANN with MLP, SVM with SMO, C4.5, CHAID and CRT from the decision trees family. These constructed models are compared in terms of classification accuracy along with misclassification rates and AUROC (area under receiver operating characteristic curve). Classification accuracy is a straightforward method considering the ratio of true estimates, which is employed widely by practitioners. ROC curve is the plot of the true positive rate against the false positive rate. That is to say, the value of the AUROC is usually between 0,5 and 1, the value closer to 1 represents a good classification whereas diagonal line 0,5 represents the test with no discriminating power.

4.1. Bayesian models. Naive Bayes and Bayesian Network are the selected classifiers representing Bayesian models. Naive Bayes classifier is a probabilistic classifier based on Bayesian theorem, often stumbles at independence assumption whereas Bayesian Network without independence assumption overcome that block. In the study both models are tested. For Bayesian Network simple estimator is chosen as estimator and Look Ahead Hill Climbing Algorithm (LAGD Hill Climbing) selected for search algorithm due to its better classification results. Classification and AUROC figures of both classifiers for each period presented in Table 4 below.

Table 4. Prediction results for Bayesian Network and Naive Bayes

Model	Performance measures	-1	-2	-3
Bayesian Network	Classification acc. (%)	91,1	84,6	75,7
	Type I error (%)	30,0	28,1	65,5
	Type II error (%)	4,6	12,6	15,9
	AUROC	0,901	0,884	0,693
Naive Bayes	Classification acc. (%)	92,2	80,2	80,9
	Type I error (%)	33,0	37,5	41,3
	Type II error (%)	2,6	16	14,5
	AUROC	0,934	0,868	0,782

In one annual period prior to failure Naive Bayes model performs slightly better than Bayesian Network. It produces 33% Type I error and 2,6% Type II error while Bayesian model produces 30% Type I error and 4,6% Type II error. As a consequence Naive Bayes' and Bayesian Network's classification accuracy rates are 92,2% and 91,1%, respectively. The AUROC results of Naive Bayes and Bayesian Network are 0,934 and 0,901, respectively.

In two annual periods prior to failure unlike the previous period above the results are in favor of Bayesian Network. It produces 28,1% type I error and 12,6% type II error while Naive Bayes produces 37,5% type I error and 16% type II error, classification accuracy of Bayesian Network and Naive Bayes are 84,6% and 80,2% respectively. The AUROC results of Bayesian Network and Naive Bayes are 0,884 and 0,868.

In three annual periods prior to failure Naive Bayes model performs better than Bayesian Network as in one annual period prior to failure. It produces 41,3% Type I error and 14,5% Type II error while Bayesian model produces 65,5% Type I error and 15,9% Type II error. The classification accuracy of Naive Bayes and Bayesian Network amount 80,9% and 75,7%, respectively. The AUROC results of Naive Bayes and Bayesian Network are 0,782 and 0,693, respectively.

Both models reach their best results in one annual period prior to failure; moreover, when the performance evaluation is considered in the scope of Type I error and AUROC results, both models are superior to each other. In regard to Type I error Bayesian Network produces fewer error than Naive Bayes in one and two annual periods prior to failure whereas Naive Bayes is superior in three annual periods prior to failure. According to AUROC results Bayesian Network is only superior in two annual periods prior to failure while in other two periods Naive Bayes has the best results.

4.2. K-NN instance based learning. K-Nearest Neighbor classifier's distance computation parameter is set to Euclidian metric with cross-validation. Distance weight parameter is set to weight by 1/distance and 3 is the selected number of neighbors to be used in classifier for each period. Classification and AUROC results are presented in Table 5 below.

Table 5. Prediction result for k-NN

Model	Performance measures	-1	-2	-3
K-NN	Classification acc. (%)	90,5	82,4	82
	Type I error (%)	40	65,6	79,3
	Type II error (%)	3,3	7,3	5,5
	AUROC	0,912	0,703	0,685

Classification accuracy of k-NN model for one annual period prior to failure is significantly better than the results of other two periods that shows closer results. Type I error results are 40%, 65,5% and 79,3% by order of periods from closest to farther period. It should be said that the longer is the period before failure, the greater the Type I error is produced. By the same period order Type II error results are 3,3%, 7,3% and 5,5%, respectively. The calculated classification accuracy values are 90,5%, 82,4% and 82%, respectively. Revealed AUROC values of the model are 0,912, 0,703, and 0,685, respectively. As it is noticeable that AUROC values have a reverse relationship with that of Type I error, the higher is the Type I error the fewer are the AUROC values.

4.3. ANN with Multilayer Perceptron (MLP). To apply ANN in classification Multilayer Perceptron (MLP) classifier, which uses back propagation algorithm for classification, is selected for training and validation. Classification and AUROC results are presented in Table 6 below.

Table 6. Prediction results for MLP

Model	Performance measures	-1	-2	-3
ANN	Classification acc. (%)	90	85,7	76,3
	Type I error (%)	40	37,5	75,8
	Type II error (%)	4	9,3	13,1
	AUROC	0,900	0,861	0,554

This ANN model shows best performance in one annual period prior to failure, overall percentage of correct classification is 90%. Produced Type I and Type II errors are 40% and 4%, respectively. For the next period achieved classification accuracy rate is 85,7% while Type I and Type II errors are 37,5% and 9,3%, respectively. Moreover classification accuracy for three annual periods prior to failure is 76,3% which is its lowest rate. AUROC values have a significant decrease tendency by the periods goes farther. The values are by order of 0,900, 0,861 and 0,554. The AUROC value for the period -3 is closer to diagonal line in other terms the model has poor classification capacity for this period. Except Type I error, which the model reach its fewest error rate, all of the indicators tend to get deteriorate by the period goes farther.

4.4. Support Vector Machines (SVM). For model construction SVM classifier with John Platt's sequential minimal optimization algorithm is selected for training and validation process of classifier. As it is explained in the synopsis part of the WEKA platform for the classifier, this algorithm globally replaces all missing values and transforms nominal attributes into binary ones and it also normalizes all attributes by default. The preferred kernel function of the algorithm is the RBF kernel function and algorithm parameters C and γ are vary through the periods. Prediction results and parameters presented in Table 7 below.

Table 7. Prediction results for SVM

Model	Performance measures	-1	-2	-3
SVM	Classification acc. (%)	92,7	86,8	85,5
	Type I error (%)	30	53,1	58,6
	Type II error (%)	2,6	4,6	5,5
	AUROC	0,952	0,711	0,679
	C	100	150	25
	γ	0,0001	0,2	1

SVM classifier achieves the best accuracy 92,7% in one annual period prior to failure. For the other two and three annual period prior to failure shows 86,8% and 85,5% classification accuracies, respectively. Type I error production for the same period rank are 30%, 53,1% and 58,6%, respectively, while Type II error rates are 2,6%, 4,6% and 5,5%, respectively. It is remarkable that Type I and Type II error have the similar course of deterioration in both case the indicators are closer for -2 and -3 periods and the error rates of -3 period almost two fold that of -1 period. AUROC value of the model in period -1 is 0,952 then for the earlier periods the value decrease drastically to 0,711 and 0,679, respectively. All of the indicators are consistent with each other proving that the correct estimation of the distressed and non-distressed firms decreases gradually in the preceding periods.

4.5. Decision trees. Selected decision tree algorithms Quinlan's C4.5, CHAID and CRT used for model construction in this study. J48 algorithm of WEKA platform represents Quinlan's C4.5. CHAID and CRT algorithms are conducted by employing SPSS 15. Revealed results for the decision tree algorithms are presented in Table 8 below.

Table 8. Prediction results for decision trees

Model	Performance measures	-1	-2	-3
C4.5	Classification acc. (%)	87,2	82,4	80,9
	Type I error (%)	46,6	25	79,3
	Type II error (%)	6	16	6,9
	AUROC	0,732	0,804	0,570
CHAID	Classification acc. (%)	92,2	88,4	83,2
	Type I error (%)	46,6	31,2	100
	Type II error (%)	0	7,3	0
	AUROC	0,767	0,807	0,500
CRT	Classification acc. (%)	96,6	99,5	95,3
	Type I error (%)	20	0	27,5
	Type II error (%)	0	0,6	0
	AUROC	0,981	0,997	0,974

At first glance the noteworthy thing from the table above that CHAID and CRT algorithms seem superior to C4.5. But when the table is examined carefully it could be seen that CHAID provides insignificant results for the period -3. Surprisingly CHAID algorithm fails at this period. Interpretation of the failure comes later in this part.

In one annual period prior to failure, CRT algorithm outperforms C4.5 and CHAID. It produces 20% type I error while C4.5 and CHAID both produce 46,6% error. In contrast, produced Type II error significantly low for the classifiers; except C4.5, which produces 6% error, CHAID and CRT classify non-distressed firms without error. The overall prediction accuracy amounts to 87,2%, 92,2% and 96,6% for the classifiers C4.5, CHAID and CRT respectively. The best AUROC figure achieved by CRT at 0,981 value. The other models have significantly lower level of AUROC that amounts to 0,767 for CHAID and 0,732 for C4.5.

In the next period CRT is superior to both models and C4.5 is still worst in classification accuracy. Classification accuracy rates of the models are 99,5%, 88,4% and 82,4%, respectively. CHAID has the highest Type I error at 32,1%, this followed by C4.5 with 25% Type I error and CRT produces zero Type I error. In Type II error production CRT model has the fewer error with 0,6% and C4.5 and CHAID produced 16% and 7,3% Type II errors, respectively. C4.5 and CHAID have closer AUROC values of 0,804 and 0,807, respectively. CRT has the highest AUROC value in this period with 0,997.

In three annual period prior to failure, CTR model has the higher classification accuracy with 95,3% and the other models C4.5 and CHAID have classification accuracies slightly over 80% with 80,9% and 83,2% by order. The highest Type I error with 100% reached by CHAID and it is followed by C4.5 with 79,3% Type I error, CRT produced acceptable Type I error of 27,5%. In Type II error production CHAID and CRT have zero errors while C4.5 produces just 6,9% Type II errors. AUROC value of CHAID points out an interesting upshot with the value of 0,500 that CHAID model fails in classification of distressed and non-distressed firms for this period since this model interprets all of the distressed firms as non-distressed. C4.5 has a low AUROC value amount to 0,570 and CRT has 0,974. When Type I error production and AUROC figures considered for classification, although CHAID's shows 83,2% classification accuracy, thank to Type I error rate C4.5 outperforms CHAID with its 80,9% classification accuracy. AUROC figures approve that assertion.

Summary and conclusion

In this study it is not aimed to present or highlight a model's superiority over others. It is aimed to present the efficiency of machine learning algorithms in financial distress prediction field.

Classification accuracies along with misclassification rates and AUROC values of representative learning algorithms for each examined period are presented. 10-fold cross validation preferred for avoiding overfitting problem since all data used for training and validation processes. All of the learning algorithms used variables selected through ANOVA and cost sensitive variable election process, in other terms the variables minimizing Type I error maximizing overall classification are used in modeling.

In one annual period prior to failure, except C4.5 all of the algorithms produce more or equal than 90% classification accuracy where CRT algorithm has the highest accuracy with 96,6%, least value belongs to C4.5 with 87,2%. This situation is consistent in AUROC values too; C4.5 has 0,732 whereas other learning algorithms have more than 0,900 AUROC values.

In the next previous period prior to failure, Naive Bayes shows the least performance in classification accuracy with 80,2% while CRT reaches 99,5% accuracy in this term. The best AUROC value belongs to CRT with 0,997 and this followed by Bayesian Network by 0,884, the least AUROC value reached by k-NN algorithm with 0,703.

In three annual period prior to failure, CHAID exhibits an interesting stage that CHAID has a moderately high classification accuracy of 83,2% which is exactly equal the ratio of non-distressed firms in whole data set, on the other hand Type I error and AUROC figures indicate that this learning algorithm is the worst in this period. When all other indicators are neglected and only classification accuracy is considered then CHAID algorithm ranks in the 3rd place among 8 learning algorithms. In fact this is not the case, this study uses AUROC and misclassification figures along with classification accuracy for evaluation of models; therefore together with AUROC and Type I error evaluation proves that CHAID algorithm is not significant in this term. It could be said that solely relying on one indicator could mislead the user. The best performer is again CRT in classification accuracy and AUROC values of 95,3% and 0,974, respectively.

In all periods CRT is the absolute winner, the promising results of CRT indicate an overfitting problem as a result of using the same data for training and validation.

More importantly, in spite of the promising results of above reported learning algorithms, this study has several limitations, some of which involve the need for additional research others are absence of robust theoretical framework for selection of potential explanatory variables of financial distress and relatively small sample size of distressed firms.

To sum up, machine learning algorithms could be used along with other statistical and structural prediction model or as an alternative tool for financial distress prediction. But assessing corporate financial structure barely relying on learning algorithm outcomes could be misleading; therefore it should be underlined that the assessment should be made by collaboration of human judgment and prediction methods.

References

1. Aghaie A., A. Saeedi. Using Bayesian Networks for Bankruptcy Prediction: Empirical Evidence from Iranian Companies // International Conference on Information Management and Engineering, Icime, 2009. – pp. 450-455.
2. Aha D.W. Lazy Learning // Special Issue Editorial. Artificial Intelligence Review 11. – Dordrecht: Kluwer Academic Publishers, 1997. – pp. 7-10.
3. Altman E.I. Financial Ratios, Discriminant Analysis and The Prediction of Corporate Bankruptcy // *The Journal of Finance*, 1968. – №23. – pp. 589-609.
4. Altman E.I., G. Marco, and F. Varetto. Corporate Distress Diagnosis: Comparisons Using Linear Discriminate Analysis and Neural Networks // *Journal of Banking and Finance*, 1994. – №18 (3). – pp. 505-529.
5. Atiya A.F. Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results // *IEEE Transactions on Neural Networks*, 2001. – №12 (4). – pp. 929-935.
6. Aurira L., R.A. Moro. Support Vector Machines (SVM) as a Technique for Solvency Analysis // German Institute for Economic Research DIW Berlin, 2008. Discussion Paper 811.
7. Back B., T. Laitinen, K. Sere, M. van Wezel. Choosing Bankruptcy Predictors Using Discriminant Analysis, Logit Analysis, and Genetic Algorithms // Turku Centre for Computer Science, Technical Report, September, 1996. – №40.

8. Beaver W.H. Financial Ratios as Predictors of Failure, Empirical Reseach in Accounting: Selected Studies // *Journal of Accounting Research*, 1966. – №5. – pp. 179-199.
9. Boritz, J.E., D.B. Kennedy. Effectiveness of Neural Network Types for Prediction of Business Failure // *Expert Systems with Applications*, 1995. – №9 (4) – pp. 503-512.
10. Coats, P.K., Fant, L.F. Recognizing Financial Distress Patterns Using Neural Network Tool // *Financial Management*, 1993. – №22 (3). – pp. 142-155.
11. Cronan, T.P., Glorfeld, L.W. and Perry, L.G. Production System Development for Expert Systems Using a Recursive Partitioning Induction Approach: An Application to Mortgage, Commercial, and Consumer Lending // *Decision Sciences*, 1991. – №22 (4). – pp. 812-845.
12. Dereli G., F. Gürgen, N. Okay. A Neural Approach for SME's Credit Risk Analysis in Turkey // P. Perner (Ed.) Sipringer-Verlag Berlin Heidelberg, 2009. – pp. 749-759.
13. Etherige H.L., R.S. Sriram. A Comparison of The Relative Costs Of Financial Distress Models: Artificial Neural Networks, Logit and Multivariate Discriminant Analysis // *Intelligent Systems in Accounting, Finance and Management*, 1997. – №6 (3). – pp. 235-248.
14. Frydman, H., Altman E.I., Kao D.L. Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress // *Journal of Finance*, 1985. – №40. – pp. 269-291.
15. Härdle W.K., R.A. Moro, D. Schäfer. Rating Companies with Support Vector Machines // German Institute for Economic Research DIW Berlin, 2004. Discussion Paper 416.
16. Henley W.E., D.J. Hand. A K-Nearest Neighbour Classifier for Assessing Consumer Credit Risk // *The Statistician*, 1996. – №45. – pp. 77-95.
17. Hu Y-C., J. Ansell. Developing Financial Distress Prediction Models: A Study of US, Europe and Japan Retail Performance // Management School and Economics, University of Edinburgh, Scotland, 2006.
18. Joos P., K. Vanhoof, H. Ooghe, N. Sierens. Credit Classification: A Comparison of Logit Models and Decision Trees // In 10th European Conference on Machine Learning. Proceedings Notes of the Workshop on Application of Machine Learning and Data Mining in Finance, Chemnitz, Germany, 1998. – pp. 59-72.
19. Ko, L.J., E.J. Blocher, P.P. Lin. Prediction of Corporate Financial Distress: An Application of the Composite Rule Induction System // *The International Journal of Digital Accounting Research*, 2001. – №1 – pp. 69-85.
20. Kirkos E., C. Spathis, Y. Manolopoulos. Data Mining Techniques for the Detection of Fraudulent Financial Statements // *Expert Systems with Applications*, 2007. – №32 (4). – pp. 995-1003.
21. Kotsiantis S., D. Tzelepis, E. Koumanakos, V. Tampakas. Efficiency of Machine Learning Techniques in Bankruptcy Prediction // 2nd International Conference on Enterprise Systems and Accounting, ICESAcc'05, Thessaloniki, Greece, 2005. – pp. 39-49.
22. Koyuncugil A.S., N. Ogulbas. Risk Modelling by CHAID Decision Tree Algorithm // *International Conference on Computational & Experimental Engineering and Sciences*, 2009. – №11 (2). – pp. 39-46.
23. Lee, W.C. Genetic Programming Decision Tree for Bankruptcy Prediction // In: Proceedings of the 2006 Joint Conference on Information Sciences, Atlantis Press, 2006.
24. Lin, F.Y., S. McClean. The Prediction of Financial Distress Using Structured Financial Data From The Internet // *IJCSS International Journal of Computers, Science and Signal*, 2000. – №1. – pp. 43-57.
25. Lin R.H., C.L. Chuang, K.W. Huang. Developing a Business Failure Prediction Model via RST, GRA and CBR // *Expert System with Applications*, 2009. – №36 (2). – pp. 1593-1600.
26. Marais, M.L., Patell, J.M., Wolfson, M.A. The experimental design of classification models: an application of recursive partitioning and bootstrapping to commercial bank loan classifications // *Journal of Accounting Research*, 1984. – №22. – pp. 87-114.
27. Messier, W.F., Hansen, J.V. Inducing Rules for Expert System Development: An Example Using Default and Bankruptcy Data // *Management Science*, 1988. – №34 (12). – pp. 1403-1415.
28. Min, J.H., Y.C. Lee. Bankruptcy Prediction Using Support Vector Machine with Optimal Choice of Kernel Function Parameters // *Expert Systems with Applications*, 2005. – №28 (4). – pp. 603-614.
29. Odom, M.D., Sharda, R. A Neural Network Model for Bankruptcy Prediction // *International Joint Conference on Neural Network*, 1990. – №2. – pp. 163-168.
30. Ohlson J. Financial Ratios and the Probabilistic Prediction of Bankruptcy // *Journal of Accounting Research*, 1980, Spring. – pp. 109-131.
31. Park, C.S., I. Han. A Case-Based Reasoning with the Feature Weights Derived By Analytic Hierarchy Process For Bankruptcy Prediction // *Expert Systems with Applications*, 2002. – №23 (3). – pp. 255-264.
32. Qintana D., Y. Saez, A. Mochon, P. Isasi. Early Bankruptcy Prediction using ENPC // *Journal of Applied Intelligence*, 2008. – №29 (2). – pp. 157-162.
33. Quinlan J.R. Discovering Rules by Induction from large collections of examples // *Expert Systems in the Microelectronic Age*, 1979 edited by D. Michie. – pp. 168-201.
34. Sarkar, S., R.S. Siriram. Bayesian Models for Early Warning of Bank Failures // *Management Science*, 2001. – №47 (11). – pp. 1457-1475.
35. Shin K.S., T.S. Lee, H.J. Kim. An Application of Support Vector Machines in Bankruptcy Prediction Model // *Expert Systems with Applications*, 2005. – №28. – pp. 127-135.

36. Sun L., P.P. Shenoy. Using Bayesian Networks for Bankruptcy Prediction: Some Methodological Issues // *European Journal of Operational Research*, 2007. – №180 (2). – pp. 738-753.
37. Tam, K.Y., Kiang, M.Y. Managerial Applications of Neural Networks: The Case of Bank Failure Predictions // *Management Science*, 1992. – №38 (7). – pp. 926-947.
38. Turkish Bankruptcy Law, web: http://www.mevzuat.adalet.gov.tr/html/1029.html.
39. Turkish Trade Law, available at http://www.mevzuat.adalet.gov.tr/html/997.html.
40. Udo, G. Neural Network Performance on The Bankruptcy Classification Problem // *Computers and Industrial Engineering*, 1993. – №25 (1-4). – pp. 377-380.
41. Vapnik V. The Nature of Statistical Learning Theory New York: Springer Verlag, 1995.
42. Varetto, F. Genetic Algorithms Applications in The Analysis of Insolvency Risk // *Journal of Banking and Finance*, 1998. – №22 (10-11). – pp. 1421-1439.
43. Vieira A.S., J. Duarte, B. Ribeiro, J.C. Neves. Accurate Prediction of Financial Distress of Companies with Machine Learning Algorithms // M. Kolehmainen et al. (Eds.): ICANNGA 2009, LNCS 5495, Sipringer-Verlag Berlin Heidelberg, 2009, pp. 569-576.
44. Wilson, R.L., R. Sharda. Bankruptcy prediction using neural networks // *Decision Sciences*, 1994. – №11 (5). – pp. 545-557.
45. Witten I.H., E. Frank. Data Mining Practical Machine Learning Tools and Techniques, Second Edition. San Francisco - Elsevier, Morgan Kaufman Publishers, 2005.
46. Yang, Z., M. Platt, H. Platt. Probabilistic Neural Networks in Bankruptcy Prediction // *Journal of Business Research*, 1999. – №44 (2). – pp. 67-74.
47. Yip, A.Y.N. A Hybrid Case-Based Reasoning Approach to Business Failure Prediction // Design and Application of Hybrid Intelligent Systems, IOS Press, Amsterdam, The Netherlands, 2003. – pp. 371-378.
48. Zheng Q., J. Yan-Hui. Financial Distress Prediction on Decision Tree Models // IEEE, Service Operations and Logistics, Informatics, International Conference, August 2007.