Economic Modelling 29 (2012) 632–640

Contents lists available at SciVerse ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

Prediction of bank financial strength ratings: The case of Turkey

Hulisi Öğüt [a,*], M. Mete Doğanay [b], Nildağ Başak Ceylan [c], Ramazan Aktaş [a]

[a] Department of Business Administration, TOBB University of Economics and Technology, Söğütözü Caddesi No: 43, Ankara, 06560, Turkey
[b] Department of Business Administration, Çankaya University, Öğretmenler Caddesi, No: 14, Balgat, Ankara 06530, Turkey
[c] Faculty of Management, Yildirim Beyazit University, Çankırı Cad. Çiçek Sok No: 3, Ulus, Ankara, Turkey

ARTICLE INFO

Article history:
Accepted 11 January 2012

Keywords:
Rating agencies
Bank financial strength rating
Financial and operational ratios
Rating prediction
Multivariate statistical model
Data mining technique

ABSTRACT

Bank financial strength ratings have gained widespread popularity especially after the recent financial turmoil. Rating agencies were criticized because of their ratings and failure to predict the bankruptcy of the banks. Based on this observation, we investigate whether the forecast of the rating of bank's financial strength using publicly available data is consistent with those of the credit rating agency. We use the data of Turkish banks for this investigation. We take a country-specific approach because previous studies found that proxies used for environmental factors (political, economic, and financial risk of the country) did not have any explanatory power and it is hard to find international data for other important factors such as franchise value, concentration, and efficiency. We use two popular multivariate statistical techniques (multiple discriminant analysis and ordered logistic regression) to estimate a suitable model and we compare their performances with those of two mostly used data mining techniques (Support Vector Machine and Artificial Neural Network). Our results suggest that our predictions are consistent with those of Moody's financial strength rating in general.. The important factors in rating are found to be profitability (measured by return on equity), efficient use of resources, and funding the businesses and the households instead of the government that shows efficient placement of the funds.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Stakeholders such as investors and creditors usually base their financial decisions on ratings. Rating measures the financial credibility of a bank, a corporation, a government, etc. Depending on the definition of the financial credibility, there are different types of rating. For example, credibility in bond rating is defined as the ability to make interest and principal payments timely. In credit rating, the credibility is defined as the ability to fulfill financial obligations when they mature. Credibility in sovereign rating is defined as the ability of a government to fulfill its financial obligations. Our focus in this paper is on bank financial strength rating. A rating agency, such as Moody's assigns bank financial strength rating and defines it as "[the] Moody's opinion of a bank's intrinsic safety and soundness" (Moody's, 2006). Moody's states that, unlike other types of rating, bank financial strength rating does not measure the ability of a bank to make timely payments, but it measures a bank's ability to avoid default. In other words, bank financial strength rating provides information to the third parties about the financial health of a bank. Moody's assigns these ratings by designating letters between A and E, and (+) (−) signs with these letters. Moody's takes into consideration some quantitative and qualitative factors when determining these ratings. Moody's groups these factors into five broad categories: franchise value, risk positioning, regulatory environment, operating environment, and financial fundamentals. Some of these factors are general factors, which apply to all banks within an environment such as a country or a region; whereas others are specific ones, which apply to individual banks. Franchise value is defined by Moody's as "the solidarity of a bank's market standing in a given geographical market or business niche". Franchise value encompasses sub-factors, such as market share and sustainability, geographical diversification, earnings stability, earnings diversification, and vulnerability to event risk (risk that an event can destroy a bank's franchise value). Risk positioning is a measure of a bank's attitude towards risk and its ability to manage risk. This factor encompasses sub-factors such as corporate governance, controls, financial reporting transparency, credit risk concentration, liquidity management, and market risk appetite. Regulatory environment and operating environment are general factors and they are not related to individual banks. These two factors define the environment in which the bank is operating. Financial fundamentals encompass sub-factors such as profitability, liquidity, capital adequacy, efficiency, and asset quality. Moody's assesses the sub-factors and factors and assigns a rating to a bank according to a score based on the assessments.

The financial health of banks is also important for an economy as bank failures erode the confidence of the investors in the financial system and reduce credit supply, which slows down the economy and may cause a recession. It is well known that the global economic

* Corresponding author. Tel.: +90 312 292 4218; fax: +90 312 292 4104.
E-mail addresses: hogut@etu.edu.tr (H. Öğüt), mdoganay@cankaya.edu.tr (M.M. Doğanay), nildag.ceylan@ybu.edu.tr (N.B. Ceylan), raktas@etu.edu.tr (R. Aktaş).

0264-9993/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.econmod.2012.01.010

crisis of 2008 was ignited by bank failures. The economic crisis of Turkey in 2001 was also worsened by bank failures. After this crisis, many legal and regulatory measures were taken and Turkish Banking Sector was reorganized. The Turkish economy faced the 2008 global economic crisis with this reorganized and well regulated banking sector and the impact of the crisis on the Turkish economy was not so severe. Many economists attributed this light impact to a strong and healthy banking sector.

Bank financial strength ratings have gained widespread popularity especially after the recent financial turmoil. Rating agencies were criticized because of their ratings and failure to predict the bankruptcy of the banks. Having motivated by these developments and the scarcity of studies related to bank financial strength ratings in the literature; our aim in this paper is to develop models to determine the significant factors that have an impact on the bank's financial strength rating. Rather than developing alternative methodologies used by rating agencies, our purpose in this paper is to determine the model that predicts the bank financial strength ratings best, the factors that are important in determining the financial strength ratings, and the objectivity of these ratings. We define objectivity as the use of objective and publicly available data rather than subjective judgments. We do not mean that subjective judgments are not important, but we want to determine the level of objectivity. For this purpose, we used quantitative proxies for some qualitative factors that are used by Moody's. Because of this approach, our study differs from other studies that have used only accounting and financial data. In addition, environmental factors can also be important in determining the ratings. However, it is very hard to quantify environmental factors and the rater's judgment plays an important role in these factors. Other studies that have used proxies for environmental factors found that these proxies did not have any explanatory power. For this reason, we do not consider environmental factors in our research. Furthermore, we did not use cross-country data in our paper as international databases provide only financial and accounting data. However, we also want to use proxies for other factors. Since we can obtain such data for Turkish banks, we confine our sample to only Turkish banks operating in the same economic and political environment. In this way, as banks cannot control the macroeconomic and political factors, we can identify bank specific factors distinguishing the ratings of the banks. These findings also emphasize what they should do with bank specific factors in order to improve their ratings relative to the other banks operating in the same environment.

The paper proceeds as follows: The second section provides the brief literature. The data are explained in the third section. Section 4 presents and discusses the empirical results of the models. The paper is concluded in Section 5.

2. Related literature

Rating can be regarded as a classification problem as the cases (banks, firms, governments, etc.) are grouped based on their ratings. Classification models have long been applied to finance problems such as financial failure, audit reports, financial information manipulation, stock price manipulation, etc. These models have also been developed to predict ratings or used to understand the determinants of ratings. Among them, pioneering study of Fisher (1959) used ordinary regression model to determine the important factors that affect risk premiums on corporate bonds. Basically, this was not a classification study that predicted group membership, but an explanatory study that aimed to explain the important factors affecting risk premiums. One of the earliest studies on bond rating was conducted by Horrigan (1966) in which a multiple regression model was developed to predict bond ratings assigned by Moody's and Standard & Poor's. Dependent variable in this study was ratings (ratings were converted into numerical form by assigning numbers running from 1 through 9), independent variables were mostly financial ratios calculated from accounting data. West (1970) criticized Horringan's approach stating that using only financial ratios calculated from one year's accounting data might be misleading. West performed the regression analyses to predict bond ratings by using Fisher's independent variables, which were proxies for earnings variability, firm's reliability of meeting its obligations, capital structure, and marketability of its bonds, instead of financial ratios. Pogue and Soldofsky (1969) also used regression analysis but their dependent variable was dichotomous taking the values of 0 and 1. Separate regression models were estimated for "Aaa and Aa bonds", "Aa and A bonds", "A and Baa bonds", and "Aaa and Baa bonds". Authors developed models that assign firms into either rating category in each model. Piches and Mingo (1973) developed M-group multiple discriminant analysis to predict bond ratings. They also used financial data as explanatory variables, but they performed factor analysis to reduce the number of variables by grouping them into factors. Piches and Mingo (1975) in another study used quadratic discriminant analysis instead of linear discriminant analysis to predict bond ratings. Belkaoui (1980) used stepwise multiple discriminant analysis with independent variables selected based on economic rationale. Martin and Henderson (1983) used rank discriminant analysis to predict bond ratings. There are other statistical methods used to predict bond ratings. Gentry et al. (1988) used n-chotomous multivariate probit model. They used cash-based funds flow components such as inventories, current liabilities, dividends, and financial ratios as explanatory variables. Rating prediction models for other securities were also developed. Peavy and Edgar (1983) used multiple discriminant analysis to predict the ratings of the commercial papers issued by bank holding companies, Peavy and Edgar (1984) in another study used multiple discriminant analysis again to predict the ratings of commercial papers issued by industrial firms. Chandy and Duett (1990) also developed models to predict commercial paper ratings. They compared the results of three methods, which were multiple discriminant analysis, logistic regression, and a data mining approach (CART-Classification and Regression Trees). All studies used financial data as explanatory variables. Neural network models were also used to predict bond ratings in the late 1980s and early 1990s. Dutta and Shekhar's (1988) study, and Singleton and Surkan's (1990) study are the early examples in which neural networks were used to predict bond rating.

In 1995 Moody's inaugurated bank financial strength ratings. Pioneering study on this subject was performed by Poon et al. (1999). They developed an ordered multiple logistic regression model to predict bank financial strength ratings in which cases were 130 banks that came from more than 30 countries, and explanatory variables were bank specific financial data and ratios that covered profitability, asset management and risk measures. Ratings belonged to the year 1997; financial variables belonged to the year 1996. Authors performed factor analysis in order to reduce the number of variables by grouping them into factors. They also used an aggregate measure (between 0 and 100) representing political, economic, and financial risk of the country in which the bank was operating. This measure was obtained from International Country Risk Guide. Short-term debt ratings and long-term debt ratings of the banks were two other explanatory variables that were used in the model. The analyses revealed that loan provisions were the most important factor explaining bank financial strength rating, followed by risk, and then profitability. It was found that country risk measure was not a significant factor explaining the ratings. It was also found that models that included short-term and long-term debt ratings had better predictive powers. Boyacıoğlu and Kara's paper (2007) is another example that predicted Moody's bank financial strength ratings. Their dependent variable was binary. They tried to predict only D and E ratings (ignoring gradations). Only Turkish Banks were included in this study. Ratings covering the period 2001–2005 were used. Independent variables were 20 bank specific financial ratios grouped by factor analysis. Models were developed for discriminant analysis, logistic regression, and neural networks. In the hold-out sample they did not find any significant difference in the prediction power of the models.

In the 2000s different types of ratings were predicted and different models were used to predict these ratings. Bennell et al. (2006) developed artificial neural network and ordered probit models to predict sovereign credit ratings. They used macroeconomic indicators as explanatory variables and found that artificial neural network model was superior to ordered probit. Credit rating attracted the attention of the researches after the publication of Basel Accords. Researchers developed different models to predict credit ratings of the companies. Doumpos and Pasiouras (2005) developed a multicriteria classification model (a value function technique named UTADIS) to predict the ratings assigned by a regional agency in the UK. They used financial ratios as the evaluation criteria. They performed hold-out sample tests by using out-of-sample and out-of-time data (firms other than the ones in model development and for a different time period). Kumar and Bhattacharya (2006) also attempted to predict credit ratings assigned by Moody's. They also used financial ratios as classification variables and developed a full-connected and back-propagation artificial neural network model. They used an appropriate portion of companies for training and another portion for testing. Researchers widely used artificial intelligence methods in predicting different types of ratings in the 2000s. Huang et al. (2004) used back-propagation neural network method and support vector machines, which is a learning machine technique that automatically extracts knowledge from a data set, to predict credit ratings. Kim (2005) used the adaptive learning network, which is an artificial intelligence technique, to predict bond ratings by using publicly available information. Cao et al. (2006) used support vector machines in predicting the ratings of the bonds issued in the USA. Authors compared the classification accuracy of vector support machines method with those of the neural networks, ordered probit, and the logistics regression. Analyses showed that support vector machines had the best performance in predicting the bond ratings.

3. Data

For our analysis, we collect 26 ratios of the banks as independent variables and their financial strength ratings as dependent variables from 2003 to 2009. The banks that have financial strength rating from Moody's are included in our sample. We excluded investment and development banks that are non-depository institutions. Note that all ratios are not financial as some of them are proxies of qualitative data. All data are obtained from Association of Turkish Banks database. The names of these ratios are given in Table 1. We have a rationale for choosing these ratios. Ratios X16–X24 are proxies for franchise value (X16 and X17 are proxies for earnings diversification which is included in franchise value). X3, X6, X7, X25, and X26 are proxies for risk positioning (including foreign exchange risk). The others are ratios related to financial fundamentals. Since we have too many ratios as explanatory variable, we perform principal component analysis (PCA) on data. We identify 6 factors having eigenvalue score greater than 1 as it is presented in Table A.2. The descriptive statistics of these factors and ratios and the correlation between them are presented in Tables A.1 and A.3a, A.3b, A3.c respectively. Then, we perform varimax rotation technique in order to get rotated factor loadings. The principal component analysis for these rotated factor loadings can also be found in Table A.3c. We found that variables are grouped in the following factors: X18, X19, X20, X21 and X22 in the first factor, X9, X12, X13 and X14 in the second factor, X4, X16 and X17 in the third factor, *X*2, X7 and X24 in the fourth factor, X23 in the fifth factor and X3 and X15 in the sixth factor. We grouped these variables together if the correlation between variable and the factor is greater than |0.65|. When we examine the factors we see that variables grouped in the first factor are related to the franchise value. Variables grouped in the second factor are related to the profitability (return on equity) and how efficiently the bank used its resources. Variables grouped in the third factor are related to revenue structure and non-performing loans. Variables grouped in the fourth factor are related to capital adequacy and foreign exchange exposure and concentration. Variable in the fifth factor is related to deposit concentration. Variables grouped in the sixth factor are related to asset structure, especially the ratio of loans in assets and the percentage of deposits placed as loans.

Table 1
The ratios used as explanatory variables.

Total equity/total assets	*X*2
Total loans/total assets	X3
Non-performing loans/total loans	X4
Non-current assets/total assets	X5
Liquid assets/total assets	X6
Liquid assets in foreign currency/total liabilities in foreign currency	X7
Net period income/assets	X8
Net income/equity	X9
Interest revenues/interest expenses	X10
Total deposits/total assets	X11
Net interest revenue (loss)/number of branches	X12
Net interest revenue (loss)/total assets	X13
Net interest revenue (loss)/number of employees	X14
Total loans/total deposits	X15
Net interest revenue/total revenue from operations	X16
Non-interest revenue/total assets	X17
Assets/total assets of the sector	X18
Loans/total loans of the sector	X19
Deposits/total deposits of the sector	X20
Number of branches/total branches of the sector	X21
Number of employees/total number of employees of the sector	X22
Personal deposits/total deposits	X23
Foreign branches/total branches	X24
Specialized loans/total loans	X25
Assets in foreign currency/liabilities in foreign currency	X26

Table 2 provides the number of rating and their frequencies for each year. The ratings of the banks are categorical variable. In addition, there are ordered relationship between them. However, the software that we used accepts only numeric variable in the dataset. For this reason, we transformed the ratings of the banks into the numeric form and assigned 1 to the lowest rating. For the other ratings, we used increment 1 as the financial strength rating of the bank improves by one grade. We also divided the dataset into two equal parts: test and training data. In order to get homogenous split, we divide the data in each year equally into two parts. Furthermore, we try to have the homogenous distribution of financial strength ratings in each year for training and test data. Note that the data belonging to the earlier periods concentrate on low ratings while the data belonging to the later periods concentrate on higher ratings. For this reason, we did not divide the data based on years (for example earlier years for training data, later years for test data) since such a split makes training and test data more heterogeneous.

4. Results

We used MATLAB for the implementation of ANN and SVM; SAS for the implementation of MDA and ordered logistic regression classifiers. We used the same set of training and test data in both datasets

Table 2
The number of ratings and their frequencies.

Year	Number of rating	E (1)	E+ (2)	D− (3)	D (4)	D+ (5)	C− (6)
2003	11	2	2	1	1	5	
2004	13	2	2		2	7	
2005	12	1	2		2	7	
2006	15		3	2	2	8	
2007	13			1	2	5	5
2008	12				2	4	6
2009	10					4	6

across techniques in order to compare the performances of the classifiers. Six attributes (factor score) are used as input variables. The output attribute takes the value of 1 to 6 based on the rating of the banks.

4.1. Ordered logistic regression

The first multivariate statistical technique that we use is ordered logistics regression (logit) model. Logistic regression (logit) model is used for estimating the probability and group membership of independent variable by making logistic transformation of linear combination of dependent variable. We used ordered logit model rather than binary logistic regression in our paper since higher number of rating given by the credit rating agency indicates that the financial status of the bank gets better. In ordered logit model, cumulative probabilities of class membership are used to derive the non cumulative probabilities of class membership and the instance is assigned into the class having the highest probability. For an n-type ordered categorical variable, the non cumulative probabilities of class membership is defined as

$$\begin{aligned} &P(Y=1)=P(Y\leq 1)\\ &P(Y=2)=P(Y\leq 2)-P(Y\leq 1)\\ &\vdots\\ &P(Y=n)=1-P(Y\leq n-1)\end{aligned}$$

where $P(Y\leq i)=\frac{1}{1+e^{-(c_i-(a_1x_1+a_2x_2+\ldots+a_nx_n))}}$, $a_1,a_2\ldots a_n$, are the parameters and $x_1,x_2\ldots x_n$ are the inputs. Note that for each i, c_i is different however $a_1,a_2\ldots a_n$ are the same.

First, we use 6 factor scores for the determination of the classes. Table 3a shows the regression output of ordered logistics. The coefficient's p value in the first part of table shows that factor 2 and factor 6 is significantly different from 0 at 1% significance level. Positive coefficient of factor 2 implies that the probability of getting higher rating increases as the factor 2 score rises. As there is positive correlation between X9, X12, X13, X21, X14 and factor 2, we expect that as X9, X12, X13, X14 scores go up, the probability of getting higher rating increases. We also find that as the factor 6 score goes up, the probability of getting higher rating decreases. Thus, as there is negative correlation between X3, X15 and factor 6, we expect that increase in X3 and X15 score increases the probability of getting higher rating. Since only two attributes are statistically different from 0 at 1% significance level, we choose these variables (Factor 2 and Factor 6) for the prediction of classes and find the accuracy rates of logistic classifier on the test data as 60.47% (26/43). We also provide confusion matrix of test data in Table 4a. Note that cut variables in Table 3a represent the constant in the cumulative probability functions. For the sake of completeness, we also perform probit analysis using our data as well. We found the accuracy rate of probit classifier on test data as 60.47% (24/43). Since our estimation results are similar to that of ordered logistic regression, we did not report the regression output of probit model in our paper.

We also perform logistic regression analysis using variables rather than factor scores. For this purpose, we choose variables having highest absolute correlation with factors. In this way, we objectively choose input variables. Thus, X18, X12, X17, X7, X23 and X15 are used for factor 1 through factor 6 respectively. We reported logistic regression results using these variables in Table 3b. We find only X12 and X15 statistically different from 0 at 1% significance level. Using X12 and X15 as the predictors of the classes, the accuracy rates of logistic classifier on the test data are found as 62.79% (27/43) and corresponding confusion matrix is reported in Table 4b.

4.2. Multiple discriminant analysis

The second multivariate statistical model used is multiple discriminant analysis (MDA). MDA is a method for combining independent variables in linear forms for classification purpose. For this purpose, MDA generates n linear functions belonging to n classes (i.e. rating). The linear multiple discriminant model is presented below.

$$Z_i=\beta_0+\beta_1X_1+\beta_2X_2\ldots+\beta_kX_k$$

In this model, Z_i represents the class score for ith class, β_k represents the weight for input variable X_k. After N scores for each class are computed, instance is assigned to the class having the highest score. In the prediction of the ratings, we use the same input variables (factor scores and variables) with the logistic regression. The class weights considered in the model are presented in Tables 5a and 5b respectively. When the input variables are factor scores and variables, we find the accuracy of MDA classifier as 53.49% (23/43) and 65.11% (28/43) respectively. We also provide confusion matrix for the test data in Tables 6a and 6b.

The interpretation of the linear discriminant score is not straightforward as ordered logistic classifier. However, a rise in the factor 2 score increases the rating of the banks as the weight of factor 2 score goes up consistently with an increase in rating score. As there is positive correlation between X9, X12 X13 and X14 and factor 2, increases in these variables increase the probability of higher rating. For the same reasoning, increase in factor 6 score decreases the score for higher rating except class 1 and 4 functions. As the correlation between X3, X15 and factor 6 is negative, higher values of these variables increase the rating of the banks in general. The same logic applies when variables are used as input variables.

Table 3a
The output of ordered logistic regression using factor scores.

(I)				(II)			
	Coefficient	Standard error	p value		Coefficient	Standard error	p value
Factor 1	−0.0165	0.1495	0.9121				
Factor 2	1.1164	0.2918	0.0001	Factor 2	1.0969	0.2756	0.0001
Factor 3	−0.0468	0.3536	0.8946				
Factor 4	0.1843	0.1930	0.3398				
Factor 5	−0.1180	0.2459	0.6313				
Factor 6	−1.1855	0.3211	0.0002	Factor 6	−1.1897	0.3099	0.0001
Cut 1	−4.3235	0.8903	0.0000	Cut 1	−4.1008	0.8079	0.0000
Cut 2	−2.5419	0.5821	0.0000	Cut 2	−2.5222	0.5672	0.0000
Cut 3	−2.0217	0.5179	0.0001	Cut 3	−2.0302	0.5036	0.0001
Cut 4	−1.1104	0.4494	0.0135	Cut 4	−1.1323	0.4308	0.0086
Cut 5	2.8355	0.6972	0.0000	Cut 5	2.8633	0.6934	0.0000
LR test	43.8709	p value of LR	0.0000	LR test	42.2859	p value of LR	0.0000

Table 3b
The output of ordered logistic regression using variables.

	Coefficient	Standard error	p value		Coefficient	Standard error	p value
X18	7.4314	8.5994	0.3875				
X12	0.0012	0.0005	0.0087	X12	0.0014	0.0004	0.0005
X17	30.8489	41.9512	0.4621				
X7	−0.0163	0.2468	0.9473				
X23	−0.3072	3.0017	0.9185				
X15	5.2897	1.8896	0.0051	X15	4.1287	1.2989	0.0015
Cut 1	2.0815	2.0891	0.3191	Cut 1	0.6974	0.9860	0.4794
Cut 2	3.8177	2.0222	0.0590	Cut 2	2.4323	0.9641	0.0116
Cut 3	4.2805	2.0445	0.0363	Cut 3	2.8903	0.9917	0.0036
Cut 4	5.1332	2.0908	0.0141	Cut 4	3.7337	1.0378	0.0003
Cut 5	9.2162	2.6724	0.0006	Cut 5	7.6498	1.6185	0.0000
LR test	43.8709	p value of LR	0.0000	LR test	42.2859	p value of LR	0.0000

4.3. PNN model

Probabilistic neural network (PNN) is first developed and proposed by Specht (1990). There are four layers in the structure of PNN: input layer, pattern layer, summation layer and decision layer. The neurons are distributed to pattern layer by input layer. Probability density function is estimated by using multi dimensional kernels in the pattern layer. In the summation layer, posterior probability density function is computed for each class by using Bayes' rule and pattern is classified in the decision layer based on these probabilities. One of the advantages of Probabilistic Neural Network over back-propagation networks is the ability of quick learning. Thus, PNN requires less computation compared to traditional neural network models. The performance of PNN classifier depends on the smoothing factor parameter. Three-fold cross validation technique is used in order to determine the smoothing factor parameter. The cross validation technique is used to prevent over learning of training data. In an n-fold cross validation technique, n equal parts are obtained from training data at first. Then, $n-1$ parts are used in order to develop model and the remaining part is used in order to test the model. This process is repeated until all parts of data are used as test data for the model developed by the remaining part of training data. Once all parts of data are tested, the smoothing parameter with highest average classification accuracy is chosen as optimal parameter and the performance of PNN classifier is validated with this parameter on the test data.

Table 4a
The confusion matrix of logistic classifier for test data using factor scores.

Actual/predicted	1	2	3	4	5	6
1	1	0	0	0	1	0
2	0	3	0	0	1	0
3	1	0	0	0	1	0
4	0	1	0	0	5	0
5	1	0	0	0	16	2
6	0	0	0	0	4	6

Table 4b
The confusion matrix of logistic classifier for test data using variables.

Actual/predicted	1	2	3	4	5	6
1	0	1	0	0	1	0
2	0	3	0	0	1	0
3	0	1	0	0	1	0
4	0	1	0	0	5	0
5	0	1	0	0	17	1
6	0	0	0	0	3	7

We used the parameters chosen in logit classifiers. Thus, factor 2 and 6 are chosen as input attributes at first. In our dataset, we varied smoothing parameter from 0 to 2 by 0.05 interval and we found the smoothing parameter having the highest cross validation rates as 0.3. The accuracy rate of PNN classifier in test data choosing the smoothing parameter in this manner is found to be 55.81% (24/43). We also provide confusion matrix of test data in Table 7a. We repeat the same analysis using variables of X12 and X15 as the input variables. We obtained 62.79% (27/43) using the optimal smoothing parameters derived from training data and we reported corresponding confusion matrix in Table 7b.

4.4. SVM model

SVM finds optimal hyper-plane by using quadratic programming technique in order to minimize misclassification error and maximize margin between hyper-plane and nearest point. The solution procedure of this quadratic programming formulation is Lagrange multiplier technique. When optimal hyper-plane that separates one class from the other is constructed, classification decision is given by using Kernel function (Burges (1998) and Vapnik (1995)). Radial Basis Function (RBF) is chosen as a kernel for Support Vector Machine in this paper. Although there is no established procedure for determining best kernel function, the advantages of using RBF as kernel are following: (i) the performance of sigmoid kernel is similar to RBF for certain parameters (Lin and Lin, 2003). (ii) the performance of the linear kernel with parameter C is found the same as RBF kernel with parameters C'.(Keerthi and Lin, 2003). Furthermore, while it is not possible for linear kernel to nonlinearly map input space into higher dimensional feature space, RBF kernel can do this (Hsu et al., 2004).

Two parameters should be chosen for the RBF kernel function: penalty parameter for the error (C) and kernel parameter (γ). However, it is

Table 5a
The factor scores as input variables and their weights in MDA model.

Constant	−4.05	−4.25	−3.82	−2.28	−0.72	−3.99
Factor 2	−1.22	−0.99	−0.58	−0.13	0.03	1.35
Factor 6	0.53	2.19	1.31	−0.35	−0.13	−1.96

Table 5b
The variables as input variables and their weights in MDA model.

Constant	−4.2871	−3.6957	−6.1949	−7.4943	−6.2373	−13.9450
X12	−0.0004	0.0011	0.0014	0.0015	0.0017	0.0035
X15	6.9490	5.7679	8.5002	11.7261	11.6636	15.1601

Table 6a
The confusion matrix of MDA classifier for test data using factor scores.

Actual/predicted	1	2	3	4	5	6
1	0	1	0	0	1	0
2	0	2	0	0	2	0
3	0	1	0	0	1	0
4	0	1	0	0	5	0
5	0	2	0	0	15	2
6	0	0	0	0	4	6

Table 6b
The confusion matrix of MDA classifier for test data using variables.

Actual/predicted	1	2	3	4	5	6
1	1	0	0	0	1	0
2	0	2	0	0	2	0
3	0	0	0	0	2	0
4	0	1	0	0	4	1
5	0	0	0	0	18	1
6	0	0	0	0	3	7

difficult to know which of the parameter combination will perform best in the test data. For this reason, researchers propose different techniques for choosing C and γ parameters. Among these techniques, we choose grid search technique proposed by Hsu et al. (2004). In the grid search technique, parameter space is searched with the combination of (C,γ). Although this method is not so complex, it is found that computational time required for this search is not so much different from other methods (Hsu et al., 2004). For the parameters, exponential sequence of $C=\{2^{-7},2^{-5},2^{-3},2^{-1},2^{1},2^{3},2^{5},2^{7},2^{9},2^{11},2^{13}\}$ and $\gamma=\{2^{7},2^{5},2^{3},2^{1},2^{-1},2^{-3},2^{-5},2^{-7},2^{-9},2^{-11},2^{-13}\}$ is considered. Then, we evaluate the performance of every combination of C and γ parameters in the training data using 3-fold cross validation as in PNN. Once we identify the parameter pair having best cross validation, we retrained SVM with optimal parameters and performance of SVM classifier is validated on the test data. We used two different input data for the evolution of SVM classifiers as in the other methods. These are factor scores (Factor 2 and Factor 6) and variables (X12 and X15). SVM classifier achieved highest cross validation accuracies when the parameter pair is x and y and the corresponding average classification accuracies of support vector machine (SVM) classifier on the test data is found to be 55.81% (24/43) and 65.11% (28/43) when input data are from factor scores and variables respectively. We also provide confusion matrix in Tables 8a and 8b.

Table 7a
The confusion matrix of PNN classifier using factor scores.

Actual/predicted	1	2	3	4	5	6
1	0	1	0	0	1	0
2	0	2	0	0	2	0
3	0	0	0	0	2	0
4	0	1	0	0	5	0
5	0	0	0	0	17	2
6	0	0	0	0	5	5

Table 7b
The confusion matrix of PNN classifier using variables.

Actual/predicted	1	2	3	4	5	6
1	1	0	0	0	1	0
2	0	2	0	0	2	0
3	0	0	0	0	2	0
4	0	1	0	0	5	0
5	0	0	0	0	18	1
6	0	0	0	0	4	6

Table 8a
The confusion matrix of SVM classifier for test data using factor scores.

Actual/predicted	1	2	3	4	5	6
1	0	1	0	0	1	0
2	1	2	0	0	1	0
3	0	1	0	0	1	0
4	0	1	0	0	5	0
5	0	2	0	0	15	2
6	0	0	0	0	3	7

Table 8b
The confusion matrix of SVM classifier for test data using variables.

Actual/predicted	1	2	3	4	5	6
1	1	0	0	0	1	0
2	0	2	0	0	2	0
3	0	0	0	0	2	0
4	0	2	0	0	3	1
5	0	0	0	0	18	1
6	0	0	0	0	3	7

4.5. Discussion of the results

We discuss the results from two views, one of which is the methodological view another one is the financial view. When we compare classifier matrices in terms of hit rate for each rating, we can say that the number of correct classification for each rating is similar to each other although there are minor differences. This shows the robustness of the performance of our classifiers. Furthermore, we reported the classifiers and their performances on the test data in Table 9 for comparison purpose. Based on the results of the analyses, we found that the accuracy rates are highest in ordered logistic regression when factor scores are used as input variables while multiple discriminant analysis and Support Vector Machine achieved highest accuracy rates when variables are used as input variables. We also observe that accuracy rates of all classifiers are higher when variables rather than factor scores are used as input. In addition, the accuracy rates of classifiers do not differ so much when variables are used as input.

When we look at the results from a financial perspective, we see that increase in the variables X3, X9, X12, X13, X14, and X15 rises the probability of getting higher rating. X3 is total loans/total assets. As the amount of loans in the assets goes up, the rating of the bank increases. In Turkey, there was a special case for the banks prior to 2002, the year when major bank restructuring took place. Most of the banks placed their resources into government debt securities. Government debt securities' yields were very high and they have significantly lower default risk. But the main function of the banks is to

Table 9
The summary of accuracy rates of classifiers.

Classifier	Accuracy rates when factor scores are used as input	Accuracy rates when variables are used as input
Ordered logistic regression	60.47%	62.79%
Multiple discriminant analysis	53.49%	65.11%
Probabilistic neural network	55.81%	62.79%
Support vector machine	55.81%	65.11%

provide funds for the households and businesses. After 2002, the yield of the government debt securities have decreased gradually and as a result banks decreased government debt securities portfolios and increased their loans. Our inference is that the rating agency perceives that as a bank increases its loan portfolio, it is acting more like a commercial bank and placing its funds more efficiently. As the yields of the government debt securities decrease, placing the funds as loans increases the revenue of the bank. This result also indicates that the rating of a bank is higher if its exposure to market risk (especially the interest rate risk) due to its government debt security investments is low. Rapid increase in interest rates and as a result decrease in the market value of government debt security portfolio of a bank whose investment in these securities was very large caused the failure of this bank and created systemic risk in the late 2000. X9 is return to equity and it is the main profitability ratio. Rating agency is assigning a higher rating to those banks whose return on equity (profitability) is higher. X12, X13, X14 are efficiency ratios. When a bank uses its resources (human capital and other capital) efficiently, it is getting a higher rating. X15 is total loans/total deposits ratio. As more and more deposits, which are the main funds of Turkish banks, are placed as loans the bank is getting a higher rating. The explanation is the same as the one for X3 since the rating agency wants the banks to act like a commercial bank (fund the households and the businesses) instead of channeling the funds it acquired to the government.

5. Conclusion

The aim of this research is twofold. One of them is to forecast the ratings of the banks by using financial and operational variables; another one is to determine the variables that play an important role in assigning the ratings. For this purpose, we use two popular data mining techniques (Support Vector Machine and Artificial Neural Network) to estimate a suitable model and compare their performances with those of two mostly used multivariate techniques (MDA and logit model). In forecasting the financial strength rating, the ordered logistic classifier performed better compared to other classifiers when factor scores are used as input variables while multiple discriminant analysis and support vector machine achieved highest accuracy rates when variables are used as input variables. The accuracy rates of all classifiers are higher when variables rather than factor scores are used as input. We also find relevant input variables for the prediction of financial strength rating of the banks in ordered logistic regression. According to the results, the most important factors are efficiency, profitability, and the proportion of loans in the assets. Rating agency is assigning a higher rating to those banks that generate high net income for its shareholders, uses its resources efficiently, and channeling its funds as loans to the households and the businesses. According to our inference, rating agencies find it less profitable and risky for the banks to place a high proportion of their funds (mainly the deposits) to government debt securities. These results may guide banks in order to get higher ratings and become better in terms of financial strength. We want to stress once again that environmental variables such as political and economic factors also play an important role in determining the ratings. But the banks cannot control the environmental factors. They can only control the bank specific factors and this research highlights what they should do with bank specific factors in order to improve their ratings relative to the other banks operating in the same environment.

In our paper, we only used data from Turkish banks since we can use proxies for efficiency and franchise value for these banks. Furthermore, although it is possible to find financial ratios for the banks all over the world, it is very difficult to find ratios such as net interest revenue/number of branches, net interest revenue/number of employees that are found to be important explanatory variables in our analyses. Also, it is our inference that the rating agencies take into account the country-specific factors when they assign a rating. For this reason, we believe that country-specific research provides more insight. For example, more banks financed the government instead of households and businesses in Turkey prior to 2002. According to our analysis, these types of banks cannot get higher ratings recently. We must note that we cannot find suitable proxies for some qualitative factors. So, the judgment of the raters also plays an important role in determining the ratings. Even with these restrictions, we believe that the performances of classifiers are quite high (up to 65% prediction accuracy) when we consider the dependent variable takes six different values. Thus, our results suggest that our predictions are consistent with those of Moody's financial strength rating in general.

Appendix A

Table A.1
The descriptive statistics of variable.

Variable	Mean	Median	Minimum	Maximum	Standard deviation
Rating	4.4302	5	1	6	1.4434
X2	0.1128	0.1165	−0.4320	0.2020	0.0818
X3	0.4797	0.4975	0.1240	0.7520	0.1516
X4	0.0290	0.0180	0.0000	0.2890	0.0400
X5	0.0603	0.0415	0.0100	0.2240	0.0474
X6	0.3406	0.3180	0.1090	0.7350	0.1292
X7	5.8670	0.6070	0.1890	432.7500	46.5810
X8	0.0094	0.0080	−0.0300	0.0460	0.0085
X9	0.0746	0.0685	−0.1550	0.2440	0.0553
X10	1.7500	1.7205	0.4602	3.8376	0.4484
X11	0.6527	0.6453	0.4391	1.0377	0.0875
X12	1289.9000	954.2500	−2130.1000	4943.4000	1146.6000
X13	0.0200	0.0179	−0.2128	0.0699	0.0324
X14	63.5650	41.0000	−86.3910	243.7800	58.3850
X15	0.7652	0.7695	0.1084	1.5598	0.3031
X16	0.5673	0.6085	−1.0080	0.9730	0.2643
X17	0.0144	0.0095	−0.0010	0.1480	0.0170
X18	0.0647	0.0362	0.0020	0.1880	0.0525
X19	0.0651	0.0505	0.0020	0.1460	0.0433
X20	0.0669	0.0360	0.0020	0.2320	0.0577
X21	0.0636	0.0492	0.0044	0.1938	0.0474
X22	0.0624	0.0538	0.0037	0.1848	0.0416
X23	0.3011	0.3006	0.0510	0.5409	0.1184
X24	0.0102	0.0084	0.0000	0.0523	0.0092
X25	0.0403	0.0000	0.0000	0.5985	0.1174
X26	0.8383	0.8905	0.1760	1.0570	0.1709
X27	0.2404	0.2290	0.2015	0.2919	0.0292
X28	3.6163	3.5000	3.1000	4.6000	0.4920
FR1	0.0000	−1.2153	−2.8402	3.8931	2.1254
FR2	0.0000	−0.1363	−5.3089	3.7592	1.7356
FR3	0.0000	0.5263	−12.9510	3.4412	1.7531
FR4	0.0000	0.1487	−11.9400	1.2229	1.4227
FR5	0.0000	−0.0763	−2.7697	3.1677	1.2935
FR6	0.0000	−0.2815	−2.4330	3.0908	1.2212

Table A.2
The principal component analysis of the variables.

Factors	Eigen value	% Explained	% Cumulated
1	6.625131	26.50%	26.50%
2	5.884382	23.54%	50.04%
3	2.971074	11.88%	61.92%
4	2.698196	10.79%	72.72%
5	1.416465	5.67%	78.38%
6	1.066103	4.26%	82.65%
7	0.86506	3.46%	86.11%

Table A.3a
The correlation between variables.

	Rating	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
Rating	1												
X2	0.38	1											
X3	0.61	0.36	1										
X4	−0.2	−0.53	−0.2	1									
X5	−0.41	0.08	−0.32	0.16	1								
X6	−0.22	−0.38	−0.55	0.16	0.05	1							
X7	−0.26	−0.73	−0.26	0.17	−0.01	0.34	1						
X8	0.34	0.17	0.19	0.37	−0.16	−0.09	−0.51	1					
X9	0.44	0.21	0.18	−0.17	−0.37	−0.23	−0.01	0.38	1				
X10	0.5	0.51	0.48	−0.37	−0.3	−0.26	−0.27	0.25	0.33	1			
X11	−0.53	−0.67	−0.59	0.34	0.02	0.34	0.48	−0.24	−0.17	−0.49	1		
X12	0.51	0.35	0.34	−0.06	−0.33	−0.25	−0.22	0.48	0.78	0.42	−0.34	1	
X13	0.3	0.38	0.34	−0.19	−0.36	−0.17	−0.2	0.26	0.46	0.53	−0.3	0.57	1

Table A.3b
The correlation between variables.

	X14	X15	X16	X17	X18	X19	X20	X21	X22	X23	X24	X25	X26
Rating	0.49	0.61	0.22	−0.13	−0.15	0.02	−0.22	−0.15	−0.17	0.11	−0.17	−0.4	−0.06
X2	0.33	0.37	0.21	−0.35	0.11	0.22	0.05	0.11	0.1	−0.04	−0.57	−0.22	0.46
X3	0.28	0.93	0.15	−0.17	−0.57	−0.34	−0.63	−0.55	−0.53	0.1	−0.22	−0.54	−0.3
X4	−0.02	−0.23	−0.64	0.74	−0.03	−0.07	−0.01	−0.01	0.03	0.02	0.3	0.11	−0.45
X5	−0.34	−0.31	−0.45	0.27	0.14	0.31	0.09	0.09	0.18	−0.4	0.17	−0.21	0.14
X6	−0.2	−0.48	−0.16	0.2	0.27	0.21	0.26	0.22	0.18	−0.18	0.28	0.08	−0.03
X7	−0.21	−0.24	0.17	−0.1	−0.1	−0.15	−0.07	−0.08	−0.08	0.05	0.5	0.12	−0.38
X8	0.51	0.2	−0.33	0.6	0.12	0.12	0.1	0.12	0.11	0.11	−0.11	−0.06	−0.19
X9	0.78	0.22	0.4	−0.25	0.26	0.21	0.26	0.25	0.25	0.23	−0.29	0.11	0.08
X10	0.38	0.49	0.45	−0.27	−0.2	−0.14	−0.23	−0.16	−0.19	−0.09	−0.17	−0.2	−0.01
X11	−0.28	−0.72	−0.12	0.2	0.16	−0.08	0.28	0.22	0.17	0.07	0.49	0.51	−0.01
X12	0.98	0.32	0.27	−0.16	0.24	0.29	0.19	0.16	0.19	0.18	−0.38	−0.08	0.11
X13	0.55	0.33	0.47	−0.24	−0.04	−0.07	−0.05	0.05	0.01	0.18	−0.42	0	−0.03
X14	1	0.27	0.23	−0.11	0.27	0.3	0.23	0.22	0.22	0.24	−0.37	−0.04	0.12
X15		1	0.18	−0.19	−0.52	−0.31	−0.58	−0.51	−0.48	0.08	−0.28	−0.5	−0.33
X16			1	−0.79	−0.01	−0.14	0.01	0.06	−0.01	0.19	−0.21	0.16	0.16
X17				1	−0.05	−0.02	−0.06	−0.06	−0.04	−0.17	0.35	−0.1	−0.36
X18					1	0.87	0.98	0.94	0.96	0.17	0	0.47	0.57
X19						1	0.79	0.72	0.79	−0.05	0.02	0.04	0.57
X20							1	0.95	0.96	0.21	0.02	0.59	0.56
X21								1	0.97	0.34	0	0.6	0.46
X22									1	0.26	0.03	0.56	0.47
X23										1	−0.12	0.39	−0.19
X24											1	0.03	−0.2
X25												1	0.15
X26													1

Table A.3c
The correlation between variables.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Rating	−0.14	0.52	0.04	0.15	−0.05	−0.45
X2	0.15	0.19	0.26	0.8	−0.16	−0.33
X3	−0.52	0.30	0.06	0.2	0.09	−0.69
X4	−0.04	−0.05	−0.82	−0.31	0.11	0.14
X5	0.3	−0.54	−0.2	0.01	−0.44	−0.17
X6	0.17	−0.05	−0.09	−0.36	−0.36	0.53
X7	−0.09	−0.05	0.22	−0.92	0.04	0.13
X8	0.06	0.51	−0.72	0.31	0.04	−0.06
X9	0.24	0.81	0.11	−0.04	0.15	−0.15
X10	−0.27	0.58	0.25	0.34	−0.2	−0.11
X11	0.11	−0.23	−0.1	−0.5	0.19	0.64
X12	0.22	0.87	−0.04	0.11	0.02	−0.27
X13	−0.13	0.71	0.17	0.3	0.08	0.06
X14	0.25	0.87	−0.09	0.1	0.05	−0.22
X15	−0.48	0.31	0.08	0.21	0.08	−0.7
X16	−0.08	0.42	0.82	−0.01	0.15	0.09
X17	−0.06	−0.12	−0.93	−0.08	−0.14	0.13
X18	0.97	0.10	0	0.02	0.05	0.17
X19	0.91	0.08	−0.04	0	−0.26	−0.16
X20	0.94	0.08	0.01	0.01	0.14	0.28
X21	0.89	0.09	0.01	0.05	0.24	0.24
X22	0.94	0.07	−0.03	0.02	0.18	0.15
X23	0.14	0.15	0.02	−0.06	0.86	−0.11
X24	0.02	−0.25	−0.19	−0.66	−0.15	0.1
X25	0.36	−0.02	0.07	−0.01	0.6	0.6
X26	0.6	−0.08	0.38	0.41	−0.25	0.12
Factor 1	1	0	0	0	0	0
Factor 2		1	0	0	0	0
Factor 3			1	0	0	0
Factor 4				1	0	0
Factor 5					1	0
Factor 6						1

Table A.4
The name of the banks in our data.

BANKS
Akbank
Anadolu Bank
Denizbank
Dışbank
Finansbank
Fortisbank
Garanti
HSBC
İsbank
Koçbank
Oyakbank
TEB
Pamukbank
Tekfenbank
Vakıfbank
YKB
Ziraat Bank

References

Belkaoui, Ahmed, 1980. Industrial bond ratings: a new look. Financial Management 9, 44–51.

Bennell, Julia A., Crabbe, D., Thomas, S., ap Gwilym, O., 2006. Modeling sovereign credit ratings: neural networks versus ordered probit. Expert Systems with Applications 30, 415–425.

Boyacıoğlu, Melek Acar, Kara, Y., 2007. Türk Bankacılık Sektöründe Finansal Güç Derecelerinin Tahmininde Yapay Sinir Ağları ve Çok Değişkenli İstatistiksel Analiz Tekniklerinin Performanslarının Karşılaştırılması. Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi 22, 197–217 (published in Turkish).

Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery 2, 121–167.

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

Chandy, P.R., Duett, E.H., 1990. Commercial paper rating models. Quarterly Journal of Business and Economics 29, 79–101.

Doumpos, Michael, Pasiouras, F., 2005. Developing and testing models for replicating credit ratings: a multicriteria approach. Computational Economics 25, 327–341.

Dutta, Soumitra, Shekhar, S., 1988. Bond rating: a non-conservative application of neural networks. Proceedings of the IEEE International Conference on Neural Networks.

Fisher, Lawrence, 1959. Determinants of risk premiums on corporate bonds. Journal of Political Economy 67, 217–237.

Gentry, James A., Whitford, D.T., Newbold, P., 1988. Predicting industrial bond ratings with a probit model and funds flow components. The Financial Review 23, 269–286.

Horrigan, James O., 1966. The determination of long-term credit standing with financial ratios. Journal of Accounting Research 4, 44–62.

Hsu, C.-W., Chang, C.-C., Lin, C.-J., 2004. A practical guide to support vector classification. Technical Report, Department of Computer Science and Information Engineering, National Taiwan University.

Huang, Zan, Chen, H., Hsu, C., Chen, W., Wu, S., 2004. Credit rating analysis with support vector machines and neural networks: a market comparative study. Decision Support Systems 37, 543–558.

Keerthi, S.S., Lin, C.-J., 2003. Asymptotic behaviors of support vector machines with Gaussian kernel. Neural Computation 15, 1667–1689.

Kim, Kee S., 2005. Predicting bond ratings using publicly available information. Expert Systems with Applications 29, 75–81.

Kumar, Kuldeep, Bhattacharya, S., 2006. Artifical neural network vs linear discriminant analysis in credit ratings forecast: a comparative study of prediction performances. Review of Accounting and Finance 5, 216–227.

Lin, H.-T., Lin, C.-J., 2003. A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. Technical report, Department of Computer Science, National Taiwan University.

Martin, Linda J., Henderson, G.V., 1983. On bond ratings and pension obligations: a note. Journal of Financial and Quantitative Analysis 18, 463–470.

Moody's, 2006. Bank Financial Strength Ratings: Revised Methodology.

Peavy, John V., Edgar, S.M., 1983. A multiple discriminant analysis of BHC commercial paper ratings. Journal of Banking and Finance 7, 161–173.

Peavy, John V., Edgar, S.M., 1984. An expanded commercial paper rating scale: classification of industrial issuers. Journal of Business Finance and Accounting 11, 397–407.

Piches, George A., Mingo, K.A., 1973. A multivariate anaysis of ındustrial bond ratings. Journal of Finance 28, 1–18.

Piches, George A., Mingo, K.A., 1975. A note on the role of subordination in determining industrial bond ratings. Journal of Finance 30, 201–206.

Pogue, Thomas F., Soldofsky, R.M., 1969. What is in a bond rating. Journal of Financial and Quantitative Analysis 4, 201–228.

Poon, Winnie P.H., Firth, M., Fung, H.-G., 1999. A multivariate analysis of the determinants of Moody's bank financial strenght ratings. Journal of International Financial Markets, Institutions and Money 9, 267–283.

Singleton, J.C., Surkan, A.J.S., 1990. Neural network for bond rating improved by multiple hidden layers. Proceedings of the IEEE International Conference on Neural Networks.

Specht, D., 1990. Probabilistic neural networks. Neural Networks 3, 109–118.

Vapnik, V., 1995. The Nature of Statistical Learning Theory. Springer-Verlag, New York, NY.

West, Richard R., 1970. An alternative approach to predicting corporate bond ratings. Journal of Accounting Research 8, 118–125.