COMPARING THE BANK FAILURE PREDICTION PERFORMANCE OF NEURAL NETWORKS AND SUPPORT VECTOR MACHINES: THE TURKISH CASE

Fatih Ecer [a]

[a] Assistant Professor, Ph. D, Afyon Kocatepe University, Faculty of Economics and Administrative Sciences, Department of International Trade and Finance, Afyonkarahisar, Turkey, fecer@aku.edu.tr.

ABSTRACT

Experience from the banking crises during the past two decades suggest that advanced prediction models are needed for helping prevent bank failures. This paper compares the ability of artificial neural networks and support vector machines in predicting bank failures. Although artificial neural networks have widely been applied complex problems in business, the literature utilizing support vector machines is relatively narrow and their capability for predicting bank failures is not very familiar. In this paper, these two intelligent techniques are applied to a dataset of Turkish commercial banks. Empirical findings show that although the prediction performance of the two models can be considered as satisfactory, neural networks show slightly better predictive ability than support vector machines. In addition, different types of error from each model also indicate that neural network models are better predictors.

I. INTRODUCTION

Banking is an activity that influences all the other activities and the economy as a whole. The risk of a banking crisis suggests an important need of identifying banks with potential problems even before they face severe liquidity or solvency crises. Many reasons have been denoted for bank failures including poor management practices, expanded risk-taking, inadequate accounting standards, interest rate volatility, pervasive internal control weaknesses, and increased competition from nondepository institutions (Miletić, 2008; Alam, Booth, Thordason, 2000). The last crisis has demonstrated that banks play very important role in the economy. In contrast to past crises, the current crisis began in developed countries and their economies have been influenced adversely. Unemployment has increased considerably, investments and consumption have decreased and all the governments are looking at possible ways to exit the crisis. In such cases, early warning systems help to monitor banks and prevent similar problems. Thus, the development of early warning systems is very important for regulators and policymakers. If successful, the governments could take actions to protect their markets from financial crisis, by avoiding crises in the banking sector and poor banking supervision (Ioannidis, Pasiouras, Zopounidis, 2010; Yim and Mitchell, 2004).

In this paper, Turkish banking failures investigated. The Turkish economy is important to the world since Turkey has the world's 15th largest GDP-PPP and 17th largest nominal GDP. Nowadays, the Turkish banking sector is among the strongest in Central Asia, the Middle East, and East Europe. In 2010, five Turkish banks were listed in the Forbes Global 2000 which is an annual ranking of the top 2000 public companies in the world. Eichengreen (2002) captures attention to threats coming from Turkey and Argentina. He also argues that the problems of these countries have not received sufficient attention and they must be addressed to safeguard global financial stability. Hence, we have selected Turkey as the application domain and the proposed models for failure prediction are tested for the Turkish banking sector.

Turkey has a larger number of financial crises than many of the countries in the world. The Turkish banking sector was severely tested in the 1994 financial crisis and was recovered rapidly by the Turkish government. However, the Asian crises together with the two ruinous earthquakes had a negative impact on the Turkish economy and the banking sector at the end of 1990s. A November 2000 crisis led to a significant erosion of the capital base of the banking sector and exposed further the fragility of the system. The political uncertainties, the loss of credibility of the exchange rate regime, and the abolition of the exchange rate peg in February 2001 further hit the banking sector. During the November 2000 and February 2001 crises, some banks failed financially.

Consequently, some banks ended their operations and some banks were taken over by the Savings Deposit Insurance Fund. The last crises affected 25 percent of the banks in Turkey and the total cost of them was estimated to be 50 billion dollars (Boyacioglu, Kara, Baykan, 2009; Ozkan-Gunay and Ozkan, 2007).

The study of bank failures is interesting because if examiners can detect problems early, regulatory actions can be taken either to minimize the cost to the governments or to prevent a bank from failing. It is therefore desirable to explore new prediction models and to provide early warnings to regulatory agencies (Thomson, 1991; Tam and Kiang, 1992). Since 1990s, artificial neural networks (ANNs) have widely been used to solve complex problems and have proven quite successful in many business applications. They can be an efficient tool to reveal critical issues in business decisions (Wu and Wang, 2000). Recently, support vector machines (SVMs) have applied complex problems in business. The SVMs have been used to financial applications such as bankruptcy prediction, credit ratings, and the detection of insurance claim fraud (Kumar and

Ravi, 2007). However, especially in the banking failure, the literature exploiting this approach is exiguous.

The purpose of this paper is to compare the prediction performances of ANNs and SVMs for predicting Turkish banking failures. For this purpose, we implemented ANOVA test on 36 financial ratios. 22 out of 36 financial ratios were eliminated by the ANOVA and the rest 14 financial ratios were thought to be more useful in making a difference between the failed and healthy banks. Then, the data set was divided into training and validation sets. The training set and the validation set were created using the banks for one year prior to failure and two years prior to failure respectively. Both the two sets consisted of randomly chosen 34 banks in which 17 banks were failed and 17 banks were healthy. In addition, the output variable of two models is the status of banks failed or healthy. Finally, we applied different ANN models and SVMs to bank failure problem in the Turkish case and presented a comparison of the prediction performance of these models.

The remainder of this paper is organized as follows: Section 2 positions the survey within the existing literature on bankruptcy forecasting studies. Section 3 briefly reviews some prediction models used in this paper. Section 4 discusses the research design including sampling and variables. Experimental results are presented in Section 5. Discussions are presented in Section 6. Finally, Section 7 summarizes our conclusions.

II. RELATED STUDIES

In literature, ANNs and SVMs have been applied to solve bankruptcy problems. Both of these models are categorized into artificial intelligent models and provide examiners with more precise predictions. However, the literature on comparing these models for bank failure problems is relatively narrow in the Turkish case.

Since the early 1990's, much of the research on ANNs have focused on finance problems, with special attention to bankruptcy prediction (Du Jardin, 2010; Paliwal and Kumar, 2009). Odom and Sharda (1990) are the first researchers who used ANNs to predict the failures of firms. They developed an ANN model for bankruptcy prediction and tested using financial data from various firms. The same set of data was analyzed using multivariate discriminant analysis (MDA). Results showed that ANNs could be applicable to this problem. Tam (1991) employed backpropagation neural network (BPNN) for bank bankruptcy prediction. He showed that ANNs offered better predictive accuracy than discriminant analysis (DA), factor-logistic, K-nearest neighbour (kNN), and ID3. Tam and Kiang (1992) compared the predictive accuracy performance of linear discriminant analysis (LDA), logistic regression (LR), kNN, ID3, and two ANN models on bank failure problems. Their results demonstrated that ANNs outperformed other models for one year prior training sample, whereas for two years prior training sample DA outperformed others. However, ANNs outperformed others in both the one year prior and the two years prior validation samples. Coats and Fant (1993) proposed an ANN model as an alternative method of the same ratios used by MDA. They showed ANN model outperformed MDA. Bell (1997) compared LR and ANNs in predicting bank failures. He concluded ANNs and LR performed equally well. Swicegood and Clark (2001) compared DA, ANNs, and human judgment in predicting bank failures. They found ANNs outperformed other two models. In addition, research efforts have been directed to the integration of ANN models with other soft computing tools such as genetic algorithm (Tsakonas, Dounias, Doumpos, Zopounidis, 2006), fuzzy sets (Tung, Queka, Cheng, 2004), and rough sets (Zaini, Shamsuddin, Jaaman, 2008). For further literature, readers may refer to Salchenberger, Cinar and Lash (1992), Zhang, Hu, Patuwo and Indro (1999), Lee, Booth and Alam (2005), Chen,

Marshall, Zhang and Ganesh (2006), and Kim (2011).

Recently, SVMs have gained popularity due to many attractive features and generalization performance on a wide range of problems. SVMs became a focus of interest for failure prediction and the applications of SVMs into financial failure prediction began in 2005. However, especially in the banking failure, the literature exploiting this approach is relatively narrow. For example, Boyacioglu, Kara and Baykan (2009) compared ANNs, SVMs, k-means cluster analysis, and LR to the bank failure prediction problem. They found that SVMs outperformed the majority of other approaches.

Min and Lee (2005) proposed SVMs for bankruptcy prediction. They compared the SVMs with MDA, logit, and ANNs. They concluded that the SVMs outperformed other models for the training and validation data. Wu, Tzeng, Goo and Fang (2007) suggested a model which was a genetic algorithm-based SVMs to predict bankruptcy and compared this model with that of DA, logit, probit, and ANNs. Their findings showed that the genetic algorithm-based SVMs model performed the best predictive accuracy. Chen (2011) compared some models using data collected from 200 Taiwan firms. He revealed that SVMs provided a good balance of high-accuracy short and long-term performance predictions for healthy and failed firms. Feki, Ishak and Feki (2012) classified the banks using the Bayesian and SVMs models. They found that SVMs were shown to be superior to Gaussian Bayes models. Yeh, Chi and Hsu (2010) proposed a prediction of business failure model to increase accuracy with the integration of rough set theory and SVMs. Their results showed that rough set theory with SVMs model provided better classification results than rough set theory with BPNN model.

III. PREDICTION MODELS FOR BANK FAILURES

Real-world financial data and its underlying economic processes are often nonlinear in nature. The rapid developments of computing power have allowed nonlinear models to become applicable to modeling and forecasting a host of financial relationships. ANNs are very suitable to study problems in finance with poorly defined system models and presence of nonlinear effects. They are parallelized computing systems that have the ability to learn from examples and to adapt to new conditions. In other words, ANNs do not ignore past information; instead past information's importance will be steamed up step by step as new examples are fed into the network (Aminian, Suarez, Aminian, Walz, 2006; Boyacioglu, Kara, Baykan, 2009; Tam and Kiang, 1992). They contain mathematical and algorithmic elements that mimic the biological neural networks of the human nervous system. ANNs contains two working phases, the phase of learning and that of recall. Like the synapses in the brain, the weights determine the power of the signals between the layers. As the network learning, the weights are adjusted until the error is at the minimum. The recall phase is realized by one pass using the weight obtained in the learning phase (Demyanyk and Hasan, 2010; Kauko, 2003; Boyacioglu, Kara, Baykan, 2009). There are several neural network architectures to perform various categories of tasks. In this paper, Multilayer Perceptron (MLP) and Radial Basis Function (RBF) were utilized to predict bank failures.

The MLP network was developed by Rumelhart, Hinton and Williams (1986). This network is based on a supervised learning process and a feedforward network architecture. It contains three types of layers: the input layer, one or more hidden layers, and the output layer. Any neuron of a layer is connected with another neuron of the following layer.

The neurons in the input layer receive the signals from the outside and deliver them for the neurons in the next layer. During this process, the neurons of the output layer send the information of the hidden layer(s) neurons for the outside (Kauko, 2003; Alves Portela Santos, Carneiro Affonso da Costa, dos Santos **Coelho, 2007).** Statistically, each input neuron represents one of the independent variables, whilst the output neuron(s) represent dependent variable(s) (Palmer, Montano, Sese, 2006). The training process of the MLP is usually realized with the backpropagation (BP) algorithm. The BP algorithm is based on the error correction learning rule (Alves Portela Santos, Carneiro Affonso da Costa, dos Santos **Coelho, 2007).** The error is the actual output less the output calculated by the network (n_k). In a two-layer network, the output is

$$n_k = \sum_j \left[w_{kj} \tau \left(\sum_i w_{ji} x_i \right) + w_{jC} \right]$$

(1)

w_{kj} = the weights between the hidden layer (j) and the output layer (k)

$\tau(.)$ = the nonlinear activation function of the neuron

w_{ji} = the weights between the input layer (i) and the hidden layer (j)

x_i = the input vector

w_{jC} = a bias-term, where C is a constant.

The RBF network has the same structure of layers as the MLP network, but a quite different information process in the hidden layer which increases the speed of solution. Although this network has a feedforward structure and consists of a single hidden layer and an output layer, it differs from the MLP network because it does not have weights in its hidden layer, but instead it has centers. The basic feature of this network is that all neurons in the hidden layer have locally tuned response characteristics. These neurons are fully interconnected to a number of linear neurons in the output layer (Moshiri, Cameron, Scuse, 1999; Charalambous, Charitou, Kaourou, 2000). According to Charalambous, Charitou and Kaourou (2000), the purpose of RBF networks is to transform a non-linearly separable classification problem into a linearly separable one. Once an input vector is presented to the network, the hidden unit outputs are obtained by calculating the closeness of the input vector X to the weight vector (center) of each one of the hidden units. The function used to calculate the closeness is as follows:

$$\varphi(r_i) = \exp\left[-\frac{r_i^2}{2\sigma_i^2} \right]$$

(2)

$r_i = \|X - w_i\|$: the distance between X and w_i

w_i : weight vector associated with neuron i in the hidden layer

σ_i^2 : variance associated with neuron i

X : input vector.

In order to reach the final output of the network, we have to multiply the output vector of the hidden layer by the corresponding weight vector associated with the neuron in the output layer. This weight vector is computed using the pseudoinverse method, and the given target values. Since the output neuron is linear, the actual output of the network is:

$$y = v_0 + \sum_{i=1}^{H} v_i^* \phi(r_i) \qquad (3)$$

where *H* is the number of neurons in the hidden layer.

SVMs which introduced by Vapnik (1995) are based on the structural risk minimization principle from computational learning theory and seek to minimise an upper bound of the generalization error rather than minimise the training error (Fethi and Pasiouras, 2010). The mapping function in SVMs can be either a classification function or a regression function. For classification, nonlinear kernel functions are used to transform the input data to a high dimensional feature space in which the input data become more separable compared to the original input space. Then, two parallel hyperplanes are constructed on each side of the hyperplane that separates the data by maximizing the distance between the two parallel hyperplanes. An assumption is made that the larger the margin or distance between these parallel hyperplanes the lower the generalization error of the classifier will be. The training points that are closest to the maximum margin hyperplane are called support vectors (Olson, Delen, Meng, 2012; Fethi and Pasiouras, 2010).

Let us define labeled training examples $[x_i, y_i]$, consisting of an input vector $x_i \in \Re^n$, and a class value $y_i \in -1,1$ $i = 1,...,I$. For the linearly separable case, the decision rules defined by an optimal hyperplane separating the binary decision classes are given in the following equation in terms of the support vectors (Lin, Yeh, Lee, 2011),

$$Y = sign\left\{ \sum_{i=1}^{N} \alpha_i y_i (x.x_i) + b \right\} \qquad (4)$$

where Y is the outcome, y_i is the class value of the training example x_i and represents the inner product. The vector corresponds to an input and the vectors x_i, $i = 1,..., N$, are the support vectors. In Eq. (4), b and α_i are parameters that determine the hyperplane.

As to the nonlinear separable data, it can be mapped into a high dimensional feature space with a nonlinear mapping in which we can search the optimal hyperplane. Then the problem is converted into searching the nonnegative Lagrange multipliers α_i by solving the following optimization problem,

Maximize
$$\theta(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i x_j) \qquad (5)$$

Subject to
$$\sum_{i=1}^{n} \alpha_i y_i = 0$$
$$0 \leq \alpha_i \leq C, \ i = 1,2,...,n$$

where C is a penalty parameter on the training error, which is now the upper bound on α_i is determined by the user. Hence, the final classification function is given as follows:

$$Y = sign\left\{\sum_{i=1}^{N} \alpha_i y_i K(x, x_i) + b\right\} \tag{6}$$

The function K is defined as the kernel function for generating the inner products to construct machines with different types of nonlinear decision surfaces in the input space.

Kernels have the advantage of operating in the input space, where the solution of the classification problem is a weighted sum of kernel functions evaluated at the support vectors. In this study, we conducted the experiments with four different kernels such as linear, radial basis, sigmoid, polynomial, as follows (Dash, Patra, Tripathy, 2012):

- A linear machine with kernel function

$$K(x_i, x_j) = x_i^T x_j$$

- A radial basis machine with kernel function

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad \gamma > 0$$

- A sigmoid machine with kernel function

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$$

- A polynomial machine with kernel function

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \quad \gamma > 0$$

where d is the degree of the polynomial kernel.

Linear kernel is a linear classifier and it should be used as a test of the nonlinearity in the training set, as well as a reference for the eventual classification improvement obtained with nonlinear kernels. The polynomial kernel is an efficient method for modeling nonlinear relationships. Radial basis kernel is one of the widely used kernels and usually in the Gaussian form. Finally, sigmoid kernel is related to neural networks. Figure 1 shows partitioning of any dataset with a linear kernel and with a degree two polynomial kernel (Ivanciuc, 2007).
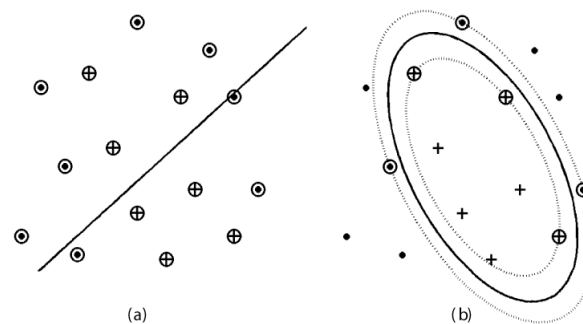

(a)   (b)

FIGURE 1. (A) LINEAR KERNEL, (B) POLYNOMIAL KERNEL WITH DEGREE 2 .

Source: Ivanciuc, 2007

IV. DESCRIPTION OF THE DATA AND VARIABLE SELECTION

The data employed in this paper were collected from The Banks Association of Turkey web site (http://www.tbb.org.tr/eng/default.aspx). The sample consists of a total of 34 Turkish commercial banks, 17 of which failed the periods 1994-2001 and contains 36 ratios available for those types of banks. Alam, Booth and Thordason (2000) indicate that it may be useful to use different time periods before bankruptcy, e.g. one year, two years, or more, before actual bankruptcy occurs.

Hence, the data sample consists of Turkish commercial banks data one year ($t-1$) and two years ($t-2$) prior to failure. 34 banks (17 failed and 17 healthy) were selected as the training set in the year $t-1$. Similarly, 34 banks (17 failed and 17 healthy) were selected in the year $t-2$ as the validation set to validate the prediction performance of each model. Hence, our comparison is based on a validation set with an equal proportion of failed and healthy banks.

In the present study, financial ratios of failed banks changed as to the date of failure. For instance, for the banks failed in 1997, $t-1$ and $t-2$ are 1996 and 1995 respectively. For the banks failed in 2001, $t-1$ and $t-2$ are 2000 and 1999 respectively. Because of most of the failures occurred between the period of 1998-2001, the year $t-1$ and $t-2$ are considered as 2000 and 1999 respectively for the healthy banks. The original classification of the failed and healthy banks is made as to the date of July 2001. Table A.1 of Appendix presents failed banks, healthy banks, and date of failure.

Firstly, we implemented ANOVA test to the 36 ratios of year $t-1$ and determined 14 ratios as the early warning indicators which have the discriminating ability for failed and healthy banks in the one year advance. Therefore, these 14 ratios can be described as early warning ratios (EWRs). Table 1 presents means and standard deviations of the ratios for the failed and healthy banks and significance tests for the equality of group means for each ratio. F statistics and their observed significance levels are shown in the last two columns. The significant level is small (<5%) for the EWRs. Hence, the EWRs are thought to be more useful in making a difference between the failed and healthy banks. In other words, they may be more suitable for predicting bank failures. In Table 1, we also calculated Wilk's lambda (λ) which is the ratio of the within groups sum of squares to the total sum of squares. λ takes the value between 0 and 1. $\lambda=1$ means all observed group means equal whereas $\lambda=0$ means within groups variability is small compared to the total variability (Canbas, Cabuk and Kilic, 2005). As can be seen in Table 1, the groups' means of the EWRs are different for failed and healthy banks.

TABLE 1. TEST EQUALITY OF GROUP MEANS FOR THE FINANCIAL RATIOS

Ratios	Failed		Healthy		Test statistics		
	Mean	SD	Mean	SD	λ	F	Sig.
Interest Expenses/Average Non-Profitable Assets	0.30	0.15	0.15	0.05	0.6666	15.989	0.000
Interest Expenses/Average Profitable Assets	0.47	0.33	0.17	0.04	0.6941	14.107	0.001
Interest Income/Interest Expenses	1.44	0.45	2.10	0.65	0.7278	11.966	0.002
(Shareholders' Equity+T.Income)/(Deposits+Non-depodit Funds)	0.03	0.19	0.23	0.15	0.7371	11.422	0.002
Total Income/Total Expenditure	0.94	0.32	1.22	0.19	0.7718	9.461	0.004
Interest Income/Average Profitable Assets	0.62	0.39	0.33	0.06	0.7745	9.308	0.005
Liquity Assets/(Deposits + Non-deposit Funds)	0.34	0.17	0.58	0.29	0.7784	9.123	0.005
Non-Interest Expenses/Total Expenses	0.30	0.11	0.41	0.11	0.7811	8.947	0.005
Interest Expenses/Total Expenses	0.70	0.11	0.59	0.11	0.7811	8.953	0.005
(Salary and Emp'ee Bene.+Res. for Retire.)/No.of Personel	0.11	0.05	0.17	0.08	0.7849	8.870	0.005
Liquid Assets/Total Assets	0.29	0.13	0.44	0.19	0.8128	7.339	0.011
(Shareholders' Equity+ T.Income)/(T.Assets+Contin.and Com.)	0.00	0.11	0.07	0.07	0.8571	5.320	0.028
(Shareholders' Equity+T.Income)/Total Assets	-0.01	0.31	0.17	0.09	0.8641	5.026	0.032
Net Working Capital/Total Assets	-0.13	0.39	0.08	0.08	0.8668	4.920	0.034
Net Income(Loss)/Average T.Assets	-0.12	0.31	0.03	0.03	0.9030	3.436	0.073
Income Before Tax / Average Total Assets	-0.11	0.32	0.03	0.03	0.9030	3.443	0.073
Non-Interest Income/Non-Interest Expenses	-0.35	0.89	0.15	0.69	0.9065	3.300	0.079
Reserve for Seniority Pay/No.of Personel (billion TL)	0.01	0.00	0.01	0.00	0.9123	3.080	0.089
Non Performing Loans/Total Loans	0.47	1.08	0.03	0.02	0.9189	2.822	0.103
Provisions except Provisons for Income Tax/Total Income	0.07	0.14	0.02	0.02	0.9294	2.422	0.129
Net Income(Loss)/Average Share-in Capital	-2.25	7.38	0.49	0.36	0.9320	2.332	0.137
Provisions including Provisons for Income Tax/Total Income	0.08	0.14	0.03	0.02	0.9363	2.211	0.147
Provision for Loan Losses/Total Loans	0.23	0.62	0.01	0.02	0.9378	2.118	0.155
Provision for Loan Losses / Total Assets	0.06	0.16	0.01	0.01	0.9411	1.935	0.174
Net Interest Income After Provision/Average T. Assets	0.04	0.22	0.11	0.06	0.9530	1.560	0.221
Fx Liquid Assets/Fx Liabilities	0.33	0.17	0.39	0.23	0.9739	0.867	0.359
Non-Interest Income/Total Income	0.81	3.56	0.02	0.22	0.9745	0.836	0.367
Interest Income/Total Income	0.19	3.56	0.98	0.22	0.9745	0.836	0.367
Net Income(Loss)/Shareholders' Equity	-0.13	1.97	0.29	0.24	0.9766	0.764	0.389
No. of Personnel / No. of Branches	21.21	6.81	22.96	8.82	0.9869	0.423	0.520
(Salaries and Emp'ee Benefits+Reserve for Retirement)/T.Assets	0.03	0.01	0.03	0.02	0.9876	0.285	0.597
Fx Assets/Fx Liabilities	0.65	0.23	0.70	0.21	0.9882	0.388	0.538
Total Loans/Total Assets	0.36	0.10	0.34	0.13	0.9934	0.269	0.608
Operational Expenses/Total Assets	0.04	0.02	0.04	0.03	0.9956	0.153	0.699
Fx Position/Shareholders' Equity	2.38	2.53	2.13	1.98	0.9967	0.104	0.749
Permanent Assets/Total Assets	0.17	0.14	0.16	0.09	0.9978	0.040	0.842

Source: Author's calculation

In this paper, the prediction performance is calculated both training and validation sets. However, when evaluating the performance of any model, its performance in validation data sets is considered as primary measures. The prediction performance is the ratio of the number of correctly classified banks to the number of incorrectly classified banks. The experiments using ANNs were performed SPSS 17.0 for Windows whereas SVMs experiments were performed in RapidMiner 5.01.

V. EMPIRICAL FINDINGS

In this section, the prediction performance of ANNs and SVMs are presented. It is determined by comparing a bank's actual membership with its predicted membership.

A. Artificial Neural Networks

The best MLP network developed to predict the bank failures is shown in Figure 2. It is composed of three layers, namely, an input layer, a hidden layer, and an output layer. Inputs for the MLP network are the 14 EWRs and output is 2 classifications, healthy or failed bank. The best MLP network contains 14 input neurons, one hidden neuron, and 2 output neurons to represent 14 explanatory variables and 2 possible outcomes. Each neuron is represented by a circle and each neuron interconnection, with its associated weight, by a line terminated by an arrow. Signals in the MLP network feedforward from left to right. The BP algorithm uses the sigmoid output function whose values range between 0 and 1. Once the network is configured, the set of initial weights is assigned at random. During the training process, the various learning rates α are examined and optimal solution is determined $\alpha = 0.4$. Besides, the momentum term μ which helps to prevent instabilities caused by a too-high learning rate is determined $\mu = 0.9$.
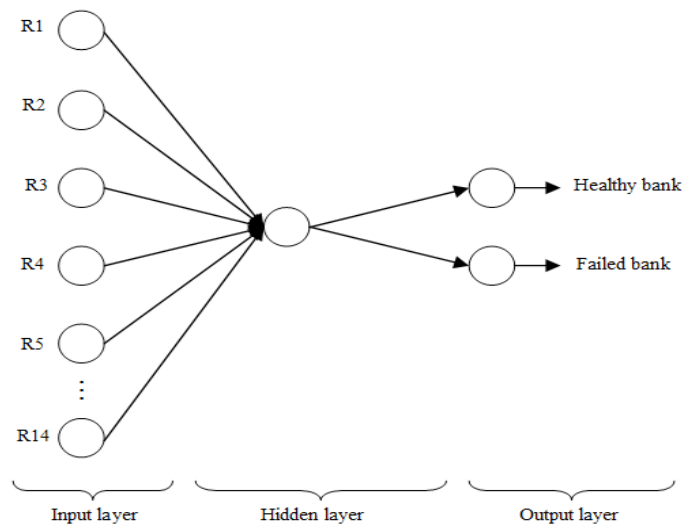

FIGURE 2. MLP ARCHITECTURE

Source: Author's calculation

The results of the MLP network are represented in Table 2 below. It accurately classifies 94.12 percent of the banks in the training data set and 97.06 percent of the banks in the validation set, two years prior to the failure. In addition, the overall error rate is 2.94 percent, with 5.88 percent type I error and 0 percent type II error in the validation set. Hence, the low type I/II errors in MLP network are relatively increased the network's reliability.

TABLE 2. PREDICTION PERFORMANCE OF MLP NETWORK

	Predicted	
Actual	Failed	Healthy
Failed	16 (94.12%)	1 (5.88%)
Healthy	0 (0%)	17 (100%)

Source: Author's calculation

The best RBF network includes 14 neurons, one hidden neuron, and 2 output neurons. Thus, it accurately classifies 88.24 percent of the banks in the training data set and 91.18 percent of the banks in the validation set. Type I/II errors are 11.76 percent and 5.88 percent respectively in the validation set. Results of RBF network are represented in Table 3 below.

TABLE 3. PREDICTION PERFORMANCE OF RBF NETWORK

	Predicted	
Actual	Failed	Healthy
Failed	15 (88.24%)	2 (11.76%)
Healthy	1 (5.88%)	16 (94.12%)

Source: Author's calculation

Finally, results of MLP and RBF networks indicate that ANNs accurately classify most of the input patterns of validation data set. As seen from the results, the performance of ANNs can be considered as satisfactory.

B. Support Vector Machines

For a comparative evaluation, we utilize four important kernels (linear, radial basis, sigmoid, and polynomial) and test on the training and the validation sets with a stopping criteria 0.001. Results for each kernel type are presented in Table 4.

Fatih Ecer

TABLE 4. PREDICTION PERFORMANCE OF SVMS

	Prediction performance	
Kernel function	Training (%)	Validation (%)
Linear	78.33	82.35
Radial basis	81.67	82.35
Sigmoid	82.35	82.35
Polynomial (1st degree)	82.35	82.35
Polynomial (2nd degree)	88.24	85.29
Polynomial (3rd degree)	85.29	79.41
Polynomial (4th degree)	79.41	79.41

Source: Author's calculation

According to Table 4, the linear, polynomial, radial basis, and sigmoid kernels have similar results that are of reasonably quality, whereas the polynomial kernel has very well classification in the validation set. In other words, SVMs yield its best performance with a second degree polynomial kernel function. SVMs accurately classify 88.24 percent of the banks in the training data set. Besides, they correctly classify 14 out of 17 healthy banks (82.35 percent) and 15 out of 17 failed banks (88.24 percent) in the validation set. Furthermore, type I/II errors are 17.65 percent and 11.76 percent respectively.

VI. DISCUSSIONS

The performance results of prediction models are presented in Table 5 and illustrated graphically in Figure 3.

TABLE 5. PERFORMANCE RESULTS FOR TRAINING AND VALIDATION OF PREDICTION MODELS

	Performance		*Errors of the validation set*		
Model	Training (%)	Validation (%)	Type I Error (%)	Type II Error (%)	Total (%)
MLP	94.12	97.06	5.88	0	2.94
RBF	88.24	91.18	11.76	5.88	8.82
SVMs	88.24	85.29	17.65	11.76	14.71

Source: Author's calculation

According to the experimental results, we conclude that prediction performance of ANNs models are more precise than SVMs in the validation set. While MLP network remains high in the ranking list, RBF network and SVMs remain second and third place respectively. In addition, type I/II errors of ANNs models are smaller than those of SVMs. MLP network remains the best classifier in terms of fewer type I, type II, and total errors. This is followed by RBF and SVMs. Besides, SVMs' best performance with a second degree polynomial kernel function and the remaining kernels has same predictions. The empirical findings reinforce the superiority of ANNs and report that ANNs are better alternatives for prediction of bank failures.
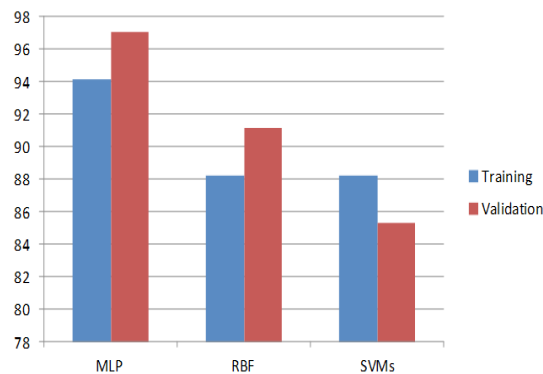

FIGURE 3. PREDICTION PERFORMANCE OF ALL MODELS

Source: Author's calculation

VII. CONCLUSIONS

In this paper, MLP and RBF networks are the most successful prediction models for bank failures. As many studies in the literature reported, the superiority of ANNs in classifying problems is proven again. However, both ANNs and SVMs are promising prediction models in identifying potentially failing banks since they achieve validation prediction performance above 85 percent. Thus, we conclude that ANNs and SVMs models adopt to solve bank failure problems and they can be useful models for bank regulators, supervisors, and others interested in early warning systems in identifying potentially failing banks. Further research could be conducted other business sectors for failure prediction.

ACKNOWLEDGEMENTS

VIII. REFERENCES

Alam, P., Booth, L. K. and T. Thordason. "The use of fuzzy clustering algorithm and self-organizing neural networks for identifying potentially failing banks: An experimental study", *Expert Systems with Applications*, 18, (2000): 185-199.

Alves Portela Santos, A., Carneiro Affonso da Costa, N. and *L.* **dos Santos** *Coelho.* "Computational intelligence approaches and linear models in case studies of forecasting exchange rates", *Expert Systems with Applications*, 33 (4), (2007): 816-823.

Aminian, F., Suarez, E. D., Aminian, M. and D. T. Walz. "Forecasting Economic Data with Neural Networks", *Computational Economics*, 28, (2006):. 71-88.

Bell, T.B. "Neural nets or the logit model? A comparison of each model's ability to predict commercial bank failures", *International Journal of Intelligent Systems in Accounting, Finance and Management*, 6, (1997): 249-264.

Boyacioglu, M.A., Kara, Y. and O.K. Baykan. "Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of savings deposit insurance fund (SDIF) transferred banks in Turkey", *Expert Systems with Applications*, 36, (2009): 3355-3366.

Canbas S., Cabuk, A. and S.B. Kilic. "Prediction of commercial bank failure via multivariate statistical analysis of financial structure: The Turkish case", *European Journal of Operational Research,* 166, (2005): 528-546.

Charalambous, C., Charitou, A. and F. Kaourou. "Comparative Analysis of Artificial Neural Network Models: Application in Bankruptcy Prediction", *Annals of Operations Research,* 99(1), (2000): 403-425.

Chen J., Marshall B.R., Zhang J. and S. Ganesh. "Financial distress prediction in China", *Review of Pacific Basin Financial Markets and Policies*, 9(2), (2006): 317-336.

Chen, M.-Y. "Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches", *Computers and Mathematics with Applications,* 62, (2011): 4514-4524.

Coats, P.K. and L.F. Fant. "Recognizing financial distress patterns using a neural network tool", *Financial Management*, 22 (3), (1993): 142-156.

Dash, S., Patra, B. and B.K. Tripathy. "A Hybrid Data Mining Technique for Improving the Classification Accuracy of Microarray Data Set", International Journal of Information Engineering and Electronic Business, 4(2), (2012): 43-50.

Demyanyk, Y. and I. Hasan. "Financial crises and bank failures: A review of prediction methods", *Omega*, 38(4), (2010): 315-324.

Du Jardin, P. "Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy", *Neurocomputing, 73*(10-12), (2010): 2047-2060.

Eichengreen, B. *Financial Crises*, Oxford University Press, New York, NY., 2002.

Feki, A., Ishak, A.B. and S. Feki. "Feature selection using Bayesian and multiclass Support Vector Machines approaches: Application to bank risk prediction", *Expert Systems with Applications*, 39, (2012): 3087-3099.

Fethi, M. D. and F. Pasiouras. "Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey", *European Journal of Operational Research*, 204 (2), (2010): 189-198.

Ioannidis, C., Pasiouras, F. and C. Zopounidis. "Assessing bank soundness with classification techniques", *Omega*, 38, (2010): 345-357.

Ivanciuc, O. "Applications of Support Vector Machines in Chemistry", *Reviews in Computational Chemistry*, 23, (2007): 291-400.

Kauko, T. "On current neural network applications involving spatial modelling of property prices", *Journal of Housing and the Built Environment*, 18(2), (2003): 159-181.

Kim, S.Y. "Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression, and multivariate discriminant analysis", *The Service Industries Journal*, 31(3), (2011): 441-468.

Kumar, P. R. and V. Ravi. "Bankruptcy prediction in banks and firms via statistical and intelligent techniques: A review", *European Journal of Operational Research*, 180(1), (2007): 1-28.

Lee, K., Booth, D. and P. Alam. "A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms", *Expert Systems with Applications*, 29(1), (2005): 1-16.

Lin F., Yeh C.C. and M.Y. Lee. "The use of hybrid manifold learning and support vector machines in the prediction of business failure", *Knowledge-based Systems*, 24 , (2011): 95-101.

Miletić, I. "Supervision of banks and the stability of the banking system", *Economic Research*, 21(3), (2008): 43-56.

Min, J.H., Y.-C. Lee. "Bankruptcy prediction using support vector machine (SVM) with optimal choice of kernel function parameters", *Expert Systems with Applications*, 28, (2005): 603-614

Moshiri, S., Cameron, N. E. and D. Scuse. "Static, Dynamic and Hybrid Neural Networks in Forecasting Inflation", *Computational Economics*, 14, (1999): 219-235.

Odom, M. and R. Sharda. "A neural network model for bankruptcy prediction", *In Proceedings of the Second IEEE International Joint Conference on Neural Networks*, Vol. II, (1990): 163-168.

Olson, D. L., Delen, D. and Y. Meng. "Comparative analysis of data mining methods for bankruptcy prediction", *Decision Support Systems*, 52, (2012): 464-473.

Ozkan-Gunay, E.N. and M. Ozkan. "Prediction of bank failures in emerging financial markets: An ANN approach", *The Journal of Risk Finance*, 8 (5), (2007): 465-480.

Paliwal, M. and U. Kumar. "Neural networks and statistical techniques: A review of applications", *Expert Systems with Applications*, 36(1), (2009): 2-17.

Palmer, A., Montano, J.J. and A. Sese. "Designing an artificial neural network for forecasting tourism time series", *Tourism Management*, 27, (2006): 781-790.

Rumelhart, D. E., Hinton, G. and Williams, R. "Learning Representation by Back-Propagating Errors", *Nature*, 323, 9, (1986): 533-536.

Salchenberger, L.M., Cinar, E.M. and N.A. Lash. "Neural networks: A new tool for predicting thrift failures", *Decision Sciences*, 23, (1992): 899-916.

Swicegood, P. and J.A. Clark. "Off-site monitoring for predicting bank under performance: A comparison of neural networks, discriminant analysis and professional human judgment", *International Journal of Intelligent Systems in Accounting, Finance and Management*, 10, (2001): 169-186.

Tam, K.Y. and M.Y. Kiang. "Managerial applications of neural networks: the case of bank failure predictions", *Management Science*, 38 (7), (1992): 926-947.

Tam, K.Y. "Neural network models and the prediction of bank bankruptcy", *Omega* 19 (5), (1991): 429-445.

The Banks Association of Turkey. http://www.tbb.org.tr/eng/default.aspx, Assessed on 12th October 2011.

Thomson, J. B. "Predicting bank failures in the 1980s", *Federal Reserve Bank of Cleveland Economic Review,* 27, (1991): 9-20.

Tsakonas A., Dounias G., Doumpos M., C. Zopounidis. "Bankruptcy prediction with neural networks by means of grammar guided genetic programming", *Expert System with Applications*, 30, (2006): 449-461.

Tung W.L., Queka C. and P. Cheng. "GenSo-EWS: A novel fuzzy based early warning system for predicting bank failures", *Neural Networks*, 17, (2004): 567-587.

Vapnik, V. N. *The nature of statistical learning theory.* New York: Springer, 1995.

Wu, C. and X.M. Wang. "A neural network approach for analyzing small business lending decisions", *Review of Quantitative Finance and Accounting,* 15, (2000): 259-276.

Wu, C. H., Tzeng, G. H., Goo, Y. J. and W. C. Fang. "A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy", *Expert Systems with Applications,* 32, (2007): 397-408.

Yeh, C., Chi, D. and M. Hsu. "A hybrid approach of DEA, rough set and support vector machines for business failure prediction", *Expert Systems with Applications*, 37, (2010): 1535-1541.

Yim, J. and H. Mitchell. "A comparison of Japanese failure models: Hybrid neural networks, logit models, and discriminant analysis", *International Journal of Asian Management*, 3, (2004): 103-120.

Zaini B.J., Shamsuddin S.M. and S.H. Jaaman. "Comparison between rough set theory and logistic regression for classification firm's performance", *Journal of Quality Measurement and Analysis*, 4(1), (2008): 141-153.

Zhang, G., Hu, M. Y., Patuwo, B. E. and D. C. Indro. "Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis", *European Journal of Operational Research*, 116, (1999): 16-32.

APPENDIX

TABLE A.1 SAMPLE OF TURKISH COMMERCIAL BANKS

Banks	Date of failure	$t-1$	$t-2$
Adabank A.Ş.	-	2000	1999
Akbank T.A.Ş.	-	2000	1999
Alternatif Bank A.Ş.	-	2000	1999
Anadolubank A.Ş.	-	2000	1999
Birleşik Türk Körfez Bank	-	2000	1999
Denizbank	-	2000	1999
Finans Bank A.Ş.	-	2000	1999
Koçbank A.Ş.	-	2000	1999
MNG Bank A.Ş.	-	2000	1999
Oyak Bank A.Ş.	-	2000	1999
Pamukbank T.A.Ş.	-	2000	1999
Tekstil Bank A.Ş.	-	2000	1999
Türk Dış Ticaret Bank A.Ş.	-	2000	1999
Türk Ekonomi Bank A.Ş.	-	2000	1999
Türkiye Garanti Bank A.Ş.	-	2000	1999
Türkiye İş Bank A.Ş.	-	2000	1999
Yapı ve Kredi Bank A.Ş.	-	2000	1999
Bank Ekspres A.Ş.	December 12, 1998	1997	1996
Bank Kapital Türk A.Ş.	October 27, 2000	1999	1998
Bayındırbank A.Ş.	July 9, 2001	2000	1999
Demirbank T.A.Ş.	December 6, 2000	1999	1998
Ege Giyim Sanayicileri Bank	July 9, 2001	2000	1999
Egebank A.Ş.	December 22,1999	1998	1997
Eskişehir Bank T.A.Ş.	December 22,1999	1998	1997
Etibank A.Ş.	October 27, 2000	1999	1998
Interbank	January 7, 1999	1998	1997
İktisat Bank T.A.Ş.	March 15, 2001	2000	1999
Kentbank A.Ş.	July 9, 2001	2000	1999
Milli Aydın Bank T.A.Ş.	July 9, 2001	2000	1999
Sitebank A.Ş.	July 9, 2001	2000	1999
Sümerbank A.Ş.	December 22,1999	1998	1997
Toprakbank	November 30, 2001	2000	1999
Türkiye Tütüncüler Bank A.Ş.	December 22,1999	1998	1997
Yurt Ticaret ve Kredi Bank	December 22,1999	1998	1997

Source: Author's calculation

Fatih Ecer

USPOREDBA PERFORMANSI NEURONSKIH MREŽA PRI PREDVIĐANJU PROPASTI BANAKA I STROJEVA S POTPORNIM VEKTORIMA: SLUČAJ TURSKE

SAŽETAK

Iskustvo stečeno u bankarskoj krizi u posljednja dva desetljeća upućuje na potrebu korištenja naprednih modela predviđanja u svrhu prevencije propasti banaka. Ovaj rad uspoređuje sposobnost umjetnih neuronskih mreža i strojeva s potpornim vektorima da predvide propast banaka. Iako se umjetne neuronske mreže često koriste za složene probleme u poslovanju, literatura koja spominje strojeve s potpornim vektorima je relativno malobrojna a njihova sposobnost predviđanja propasti banaka nije previše poznata. U ovom radu su ove dvije inteligentne tehnike primijenjene na sklop podataka turskih komercijalnih banaka. Empirijski rezultati pokazuju da iako se predviđanje dvaju modela može smatrati zadovoljavajućim, neuronske mreže pokazuju nešto bolju sposobnost predviđanja od strojeva s potpornim vektorima. Osim toga, različite vrste grešaka u svakom modelu također ukazuju na to da su modeli s neuronskim mrežama bolji prediktori.

Ključne riječi: propast banaka, ANN (umjetne neuronske mreže), SVM (strojevi s potpornim vektorima), Turska