

COMPARISON OF MACHINE LEARNING TECHNIQUES FOR ANALYZING BANKS' FINANCIAL DISTRESS

Bankaların Finansal Başarısızlıklarının İncelenmesinde Makine Öğrenme Tekniklerinin Karşılaştırılması

Gönderim Tarihi: 15.04.2016

Kabul Tarihi: 02.11.2016

Serpil ALTINIRMAK*

Çağlar KARAMAŞA**

ABSTRACT: Analyzing banks' financial distress has gained great importance due to their importance in national economy and caused sociological and economic results. Support Vector Machines (SVM) and Neural Networks (NN), known as machine learning methods, are applied for classifying banks as an early warning of financial distress. A case study which is taking thirty private equity commercial banks' five year data and financial ratios, is carried out. As a result SVM obtains better classification ratio than NNs.

Keywords: Financial Distress, Data Mining, Support Vector Machines, Financial Ratios.

ÖZ: Bankalarda meydana gelecek olan bir finansal başarısızlık sonuçlar bakımından dikkate alındığında ekonomik ve sosyolojik olarak önem arz etmektedir. Makine öğrenme tekniklerinden olan Destek Vektör Makineleri (DVM) ve Yapay Sinir Ağları (YSA) finansal başarısızlıklar konusunda erken uyarı sistemi olarak kullanılmıştır. Örnek olay olarak 30 özel sermayeli bankanın beş yıllık finansal oran verilerinden yararlanılmıştır. Yapılan analiz sonuçlarına göre destek vektör makineleri yöntemi yapay sinir ağları yöntemine göre bankalardaki finansal başarısızlıkların değerlendirilmesinde erken uyarı sistemi olarak daha iyi bir sınıflandırıcı olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: Finansal Başarısızlık, Veri Madenciliği, Destek Vektör Makineleri, Finansal Oranlar.

* Doç. Dr., Anadolu Üniversitesi/Eskişehir Meslek Yüksekokulu/Emlak ve Emlak Yönetimi Programı, sailtinirmak@anadolu.edu.tr

** Arş. Gör., Anadolu Üniversitesi/İşletme Fakültesi /İşletme Bölümü/İşletme Anabilim Dalı, ckaramasa@anadolu.edu.tr

INTRODUCTION

Firms' financial distress can be considered as essential issue in terms of sociological and economic results. Due to recession, high interest rates and tight monetary policy firms' financial risks raised. Also core internal factors such as insufficient communication, project failures, excessive growth, personnel problems, production and product errors, marketing mistakes and undercapitalization lead to financial distress (Aktaş, et. al., 2003:14). Banks' distress probability prediction has great importance because of its condition in national economy. Financial distress can be defined differently according to sectors. Financial distress can be described such as delinquency, bond yield and principal delinquency, bad cheque, assigning administration, making lose for three years in succession (Altunöz, 2013:189).

Financial distress conditions handled for banks and enterprises can be summarized as going bankrupt, delinquency, going into default, going into administration, assigning administration, bond yield and principal delinquency, halt of production due to financial crisis, bad cheque, increased non-performing loans, transferring to other bank, transferring to saving deposits insurance fund (Altunöz, 2013:190). Therefore it's important to develop an early warning system in terms of banks' bankruptcy prediction due to enable accurate and valid decisions made by shareholders, creditors and managers (Olson, et al., 2012: 464).

The rest of paper is organized as follows. Section 2 discuss relevant literature in bankruptcy prediction. Section 3 describes methodology consisting SVM and NN handled in this paper. Section 4 presents data analysis and prediction results. Section 5 gives conclusions and future recommendations.

LITERATURE REVIEW

First attempt to classify bankrupt and non-bankrupt firms is done by Beaver (1966) using financial ratios for estimating firms' distress probability during five year period. This study was followed by Altman's model (1968) based on multivariate discriminant analysis (MDA). Rose et al. (1982) asserted economic conditions as substantial reason for financial distress. Odom and Sharda (1990) compared neural networks (NN) and MDA in order to predict firms' bankrupt likelihood. As a result NN outperformed the MDA for this purpose. NNs were started to use for examining financial distress and bankruptcy with studies made by Hamer (1983), Coats and Fant (1993), Chin-Sheng, et. al., (1994), Boritz and Kennery (1995).

Studies for firms' bankruptcy prediction (Tam and Kiang, 1992; Lee, et al., 1996; Eklund, et al., 2003; Hoppszallern, 2003; Chang, et al., 2003; Lansiluoto,

et al., 2004; Kloptchenko, et al., 2004; Magnusson, et al., 2005) were made since 1990s (Koyuncugil, 2007: 3-6). Chan and Wong (2007) tried to develop data mining based early warning system for financial disrupt forecasting. Çilli and Temel (1988) used discriminant and factor analysis in order to anticipate banks' financial disruptions. Albayrak and Erkut (2005) tried to prove deficient or incorrect results of financial ratio based performance evaluation studies. Altaş and Giray (2005) utilized discriminant analysis based model by means of detecting firms under the risk of financial disruption.

Benli (2005) used logistic regression and NNs for financial distress prediction. Kılıç (2006) developed early warning model to detect financial disruptions in banking system. Boyacıoğlu, Kara and Baykan (2009) asserted the superiority of multilayer perceptron and vector quantization in bank disrupt prediction. Koyuncugil (2007) designed fuzzy data mining based early warning system to detect transaction manipulation in BIST stock exchange market. Koyuncugil and Özgülbaş (2006) determined financial profiles of small and medium sized enterprises listed in BIST via data mining methods.

METHODOLOGY

Support vector machines (SVM) ,suggested by Vapnik for solving classification and regression problems, is statistical learning and structural risk minimization principle based machine learning method (Vapnik, 1995). SVM is designed for binary classification. This method is based on seperating two classes via hyperplane in high dimensional transformation of data (Kavzoğlu and Çölkesen, 2010:74). With the introduction of SVM performances of some application areas such as regression estimation, pattern recognition, time series forecasting (Kim, 2003), medical diagnosis (Tarassenko, et al., 1995), raised also (Shin, et al., 2005:129).

Basic principle of SVM is to find maximum bound between samples by transforming nonlinear sample space into high dimensional one samples being separated linearly. In this manner SVM aims to classify new data properly after data training process.

SVMs are trained for maximizing margin so generalization ability is good even though inadequate training data. Furthermore local minimum condition is not happened. Problem can be solved by quadratic programming techniques because of SVMs' being formulated as quadratic programming problem (Abe, 2005:39).

Also SVM can classify new unobservable training datas properly. This can be considered as generalizability of SVM. SVM is a good alternative than other

techniques (neural networks, decision trees etc.) thanks to generalization (Karagülle, 2008).

Main advantages of SVM are ease of implementation and showing good performance in real world problems. SVMs are based on linear discriminant function and provable by large margin classifier. With this speciality data is separated correctly (Erastö, 2001:34).

Apart from advantages, weaknesses of SVMs are revealed as time consuming in data training process related to high dimensional problems, computation complexity and formulation redundancy (Abe, 2005:40).

Suppose a training data set $L = \{x_i, y_i\}$ ($i = 1, 2, \dots, n$) with input vector $x_i \in \mathbb{R}^n$ and target labels $y_i \in \{-1, +1\}$ in which having N training samples. According to linearly seperable training samples SVM finds an optimal hyperplane seperating binary decision classes and decision rules defined by optimal hyperplane is shown in following equation (Shin, et al., 2005 : 130):

$$Y = \text{sign} \left(\sum_{i=1}^N y_i \alpha_i (x \bullet x_i) + b \right) \quad (1)$$

where y_i is the class value of the training example (x_i), inner product is showed by (\bullet) , Y is response variable, α_i and b are hyperplane determining parameters, input vector and support vectors are represented by $x = (x_1, x_2, \dots, x_n)$ and x_i , ($i = 1, 2, \dots, N$) respectively.

For linearly seperable case optimal hyperplane separate samples with error and make margin width having biggest value between two parallel bounding planes at the opposite side of seperating plane (Hui and Sun, 2006:276).

High dimensional version of Eq. (1) is obtained in nonlinearly seperable case and showed as follows (Shin, et al., 2005:130):

$$Y = \text{sign} \left(\sum_{i=1}^N y_i \alpha_i K(x, x_i) + b \right) \quad (2)$$

where $K(x, x_i)$ as kernel function shows the inner product by different nonlinear decision surfaces in input space.

According to nonlinearly seperable training samples SVM uses nonlinear function to map input space to high dimensional feature space and nonlinear optimal seperating hyperplane having biggest margin width is found (Hui and Sun, 2006: 276).

Kernel fuctions allowing training data separate linearly in high dimensional space cause them achieving nonlinear-seperable in input space (Chen, et al.,

2009 : 10978). Three common types of kernel function namely polynomial, radial basis and sigmoid kernel functions are used in applications. Polynomial kernel function with degree of polynomial kernel (d) is showed as follows:

$$K(x, x_i) = (x \bullet x_i + 1)^d \quad (3)$$

Radial basis kernel function with bandwidth value (δ^2) is showed as follows:

$$K(x, x_i) = \exp\left(-\frac{1}{\delta^2(x - x_i)^2}\right) \quad (4)$$

Sigmoid kernel function with parameters (v and c) are showed as follows:

$$K(x, x_i) = S[(x \bullet x_i)] = \frac{1}{(1 + \exp\{v(x \bullet x_i) - c\})} \quad (5)$$

Support vector machines aims to find support vectors x_i , ($i = 1, 2, \dots, N$) and hyperplane determining parameters (α_i and b) in classification problems. Threshold values on the coefficient of α_i ranges from 0 (lower) to C (upper) for separable and non-separable cases respectively (Shin, et al., 2005 : 130).

Neural networks (NNs) comprising neuron sets, connectivity patterns, rule types (propagation, activation, transfer and learning) are parallelized computing systems which are learned from examples and adapted to new conditions (Chu, 1997; Rumelhart and McClelland, 1989). Neural networks are consisted of learning and validation stages like other machine learning methods (Neaupane and Adhikari, 2006). Multi-layer perceptron (MLP) one of the NN architecture is performed in this paper. MLP one of the supervised NNs comprised of connected feed-forward NNs with input, output and hidden layers in parallel. MLP consists minimization of error function calculated via back-propagation algorithm and trained by gradient descent method (Bose and Liang, 1996).

ANALYSIS AND RESULTS

The research data is provided by the Banks Association of Turkey and Banking Regulation and Supervision Agency and contains 30 private equity commercial banks which suffer from bankruptcy (17 cases) and non-bankruptcy (13 cases) from 1996 to 2000. Banks transferred to Savings Deposit Insurance Fund are considered as being bankrupt. Five of thirteen banks suffer from non-bankruptcy and four of seventeen banks suffer from bankruptcy are classified as foreign based. Period is handled and examined according to the multitude of banks being failed. Data set is randomly split into training (%80) and

validation/test set (%20) by cross validation in order to construct the model and test the results of the model. Cross validation method aims to prevent data loss and overcome overfitting problem. As different from other studies more financial indicators and the longer data period before bankruptcy are handled and compared in terms of two machine learning methods.

While financial ratios of banks' are considered as input variable, banks' bankruptcy condition (e.g. classifying as bankrupt or non-bankrupt) are treated as output variable. For this purpose financial ratios are divided into seven ratio groups namely liquidity, facility, profitability, size, financial structure, capital and asset condition. Financial ratios are selected according to financial literature and applications. These ratio groups consisting forty six financial ratios are given as following table:

Table 1: Liquidity Ratio Groups

Variable	Definition
X_{11}	Liquid assets/Total assets
X_{12}	Liquid assets/(deposit + non-deposit funds)
X_{13}	Foreign currency assets/ Foreign currency liabilities

Table 2: Facility Ratio Groups

Variable	Definition
X_{21}	Provisions excluding tax/ Total income
X_{22}	Provisions including tax/ Total income
X_{23}	(Personnel expense + severance)/ Number of employees
X_{24}	(Personnel expense + severance)/ Total assets
X_{25}	Severance / Number of employees

Table 3: Profitability Ratio Groups

Variable	Definition
X_{31}	Net income of the period/Average total revenue
X_{32}	Net income of the period/Average owner's equity
X_{33}	Net income of the period/Average paid-in capital
X_{34}	Income before tax/Average total assets
X_{35}	Net interest income/ Average returned assets

Table 4: Firms' Size Ratio Groups

Variable	Definition
X_{41}	Sector share – Total assets
X_{42}	Sector share – Total credits
X_{43}	Sector share – Total deposits

Table 5: Financial Structure Ratio Groups

Variable	Definition
X_{51}	Non-interest expenses/Total expenses
X_{52}	Interest expenses/Total expenses
X_{53}	Interest expenses/Average returned assets
X_{54}	Operating expenses/Total assets
X_{55}	Operating expenses/Average total assets
X_{56}	Total income/Total expenses
X_{57}	Non-interest income/Total income
X_{58}	Interest income/Total income
X_{59}	Interest income/Interest expenses
X_{510}	Net interest income/Average total assets
X_{511}	Non-interest income/Average total assets
X_{512}	Non-interest income/Non-interest expenses
X_{513}	Net income/Average capital
X_{514}	Net loss/Average capital

Table 6: Capital Ratio Groups

Variable	Definition
X_{61}	(Owner's equity + Income)/Total assets
X_{62}	(Owner's equity + Total income)/(Total assets + Non-cash credits)
X_{63}	(Owner's equity+ Total income)/(Deposit+ Non-deposit resources)
X_{64}	(Owner's equity + Total income)/Total credits
X_{65}	Net working capital/ Total assets
X_{66}	Foreign exchange position/ Owner's equity

Table 7: Asset Condition Ratio Groups

Variable	Definition
X_{71}	Total credits/ Total assets
X_{72}	Non performing loans/ Total credits
X_{73}	Total credits/ Total deposit
X_{74}	Non performing loans (gross) /Total assets
X_{75}	Non performing loans (net) /Total assets
X_{76}	Provision for non performing loans/ Total assets
X_{77}	Provision for non performing loans/ Total credits
X_{78}	Total credits/ (Owner's equity + Special provisions)
X_{79}	Foreign exchange assets/ Foreign exchange liabilities
X_{710}	Fixed assets/ Total assets

Banks' classification in terms of successful and unsuccessful are implemented by weka open source coding platform. All financial variables are included in the analysis. Failure prediction models are consisted of one (t), two $[t, (t - 1)]$, three $[t, (t - 1) (t - 2)]$, four $[t, (t - 1) (t - 2) (t - 3)]$ and five $[t, (t - 1) (t - 2) (t - 3) (t - 4)]$ years data. While parameters that represent the years before being transferred to savings deposit insurance fund vary for failed banks, it is constant ($t=2000$) for non-failure ones. From the kernel functions linear type is applied in the analysis.

Performance results of compared methodologies are showed in Table 8:

Table 8: Classification Ratios of Methods

Method	1 year data (%)	2 years data (%)	3 years data (%)	4 years data (%)	5 years data (%)
SVM	91.9	89.2	88.9	86.3	83.8
MLP	87.4	85.2	83.4	80.2	78.6

According to the Table 8, SVM predicts the financial failures of banks by using 1 year data with 91.9 percent of accuracy as compared to MLP with 87.4 percent of accuracy. Similarly SVM with 89.2 percent of accuracy performs better than MLP with 85.2 percent of accuracy by using 2 years financial data before bankruptcy. For other years SVM outperforms than MLP with having higher prediction accuracy ratio. Additionally it can be seen from the table that classification accuracy of both methods decrease as long as more years examined before bankruptcy.

CONCLUSIONS AND RECOMMENDATIONS

Financial distress defined under different concepts need to be take into the account for banks' due to their importance in national economies. Banks suffer from financial distress can go downsizing, bankrupt or transfer to other bank(s). So it is important to develop an early warning system with the purpose of predicting banks' failure and preventing from being bankrupt.

This study aims to analyze banks financial failures by machine learning methods. For this purpose two machine learning methods namely support vector machine and neural networks are compared in banks' failure prediction. Multi-layer perceptron one of the supervised neural network architecture is handled. Based on the experimental results SVM outperforms than MLP with having higher prediction accuracy ratio and serves as an alternative for financial distress prediction models. Also as more years examined before bankruptcy, classification accuracy of both methods decrease. So it can be gain more consistent and reliable results by taking less years before bankruptcy into the account. Prediction accuracy of method can be improved by handling different industry classification and period. Also the size of training set that has an effect on prediction performance can be increased.

REFERENCES

- Abe, S. (2005). *Support Vector Machines For Pattern Classification*. London: Springer.
- Aktaş, R., Doğanay, M.M. and Yıldız, B. (2003). Predicting the financial failure: A comparison of statistical methods and neural networks. *Ankara University Journal of SBF*, 58, 1-24.
- Albayrak, Y.E. and Erkut, H. (2005). Banka Performans Değerlendirmede Analitik Hiyerarşi Süreç Yaklaşımı, *İTÜ Mühendislik Dergisi/d*, 4 (6), 47-58.
- Altaş D.,and Giray, S. (2005). Mali Başarısızlığın Çok Değişkenli İstatistiksel Yöntemlerle Belirlenmesi: Tekstil Sektörü Örneği. *Marmara Üniversitesi Sosyal Bilimler Dergisi*, (2) s.13-28.
- Altman, E.I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *The Journal of Finance*, 23(4),589-609.
- Altunöz, U. (2013). Bankaların Finansal Başarısızlıklarının Yapay Sinir Ağları Modeli Çerçevesinde Tahmin Edilebilirliği, *Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 28(2), 189-217.
- Beaver, W. H.. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, 4, 71-111.
- Benli, Y.K. (2005). Bankalarda Mali Başarısızlığın Öngörülmesi Lojistik Regresyon ve Yapay Sinir Ağı Karşılaştırması. *Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dergisi*,16, 31-46.
- Boritz, E. J. and Kennery, D. (1995). Effectiveness of neural network types for prediction of business failure. *Expert Systems with Applications*, 9, 503-512.
- Bose, N.K. and Liang, P. (1996). *Neural network fundamentals with graphs, algorithms and applications*. USA: McGraw-Hill.
- Boyacıoğlu, M.A., Kara, Y. and Baykan, Ö.K. (2009). 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(2), 3355-3366.
- Chan, N.H. and Wong, H.Y. (2007). Data mining of resilience indicators. *IIE Transactions*, 39, 617-627.
- Chang, S., Chang, H., Lin, C. and Kao, S. (2003). The effect of organizational attributes on the adoption of data mining techniques in the financial service industry: An empirical study in Taiwan. *International Journal of Management*, 20(1), 497-503.

- Chen, S., Wang, W. and Zuylen, H.V. (2009). Construct support vector machine ensemble to detect traffic incident. *Expert Systems with Applications*, 36, 10976-10986.
- Chin-Sheng, H., Dorsey, R. E. and Boose, M.A. (1994). Life insurer financial distress prediction: A neural network model. *Journal of Insurance Regulation*, 13(2), 131-168.
- Chu, C.H. (1997). An improved neural network for manufacturing cell formation. *Decision Support Systems*, 20, 279-295.
- Coats, P. K. and Fant, L. F. (1993). Recognizing financial distress patterns using a neural network tool. *Financial Management*, 22 (3), 142-155.
- Çilli, H. and Temel, T. (1988). Türk Bankacılık Sistemi İçin Bir Erken Uyarı Modeli, *T.C. Merkez Bankası Kütüphanesi*, No:8804, 1-36.
- Eklund, T., Back, B., Vanharanta, H. and Visa, A. (2003). Using the self-organizing map as a visualization tool in financial benchmarking. *Information Visualization*, 2, 171- 181.
- Erastö, P. (2001). *Support Vector Machines-Backgrounds and Practice, Academic Dissertation for The Degree of Licentiate of Philosophy*. Rolf Nevanlinna Institute, Helsinki.
- Hamer, M. (1983). Failure Prediction: Sensitivity of Classification Accuracy to Alternative Statistical Methods and Variable Sets. *Journal of Accounting and Public Policy*, 2, 289-307.
- Hoppszallern, S. (2003). Healthcare bechmarking. *Hospitals & Health Networks*, 77, 37-44.
- Hui, X.F. and Sun, J. (2006). *An application of support vector machine to companies' financial distress prediction*. In the Modeling Decisions for Artificial Intelligence. Springer-Verlag, 274-282.
- Karagülle, F. (2008). *Destek Vektör Makinelerini Kullanarak Yüz Bulma*, Yüksek Lisans Tezi, Edirne: Trakya Üniversitesi.
- Kavzoğlu, T. and Çölkesen, İ. (2010). Destek Vektör Makineleri ile Uydu Görüntülerinin Sınıflandırılmasında Kernel Fonksiyonlarının Etkilerinin İncelenmesi. *Harita Dergisi*, 144, 73-82.
- Kılıç, S. (2006). Türk bankacılık sistemi için çok kriterli karar alma analizine dayalı bir erken uyarı modelinin tahmini. *ODTÜ Gelişme Dergisi*, 33(1), 117-154.
- Kim, K. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, 55(1-2), 307-319.

- Kloptchenko, A., Eklund, T., Karlsson, J., Back, B., Vanhatanta, H. and Visa, A. (2004). Combining data and text mining techniques for analysing financial reports. *Intelligent Systems in Accounting Finance and Management*, 12, 29-41.
- Koyuncugil, A. S. and OZgulbas, N. (2006). Financial profiling of SMEs: An application by data mining. The European Applied Business Research (EABR) Conference, Clute Institute for Academic Research.
- Koyuncugil, A.S. (2007). Veri Mdenciliği ve Sermaye Piyasalarına Uygulaması. *Sermaye Piyasası Kurulu Araştırma Raporu*.
- Lansiluoto, A., Eklund, T., Barbro, B., Vanharanta, H. and Visa, A. (2004). Industry-specific cycles and companies' financial performance comparison using self-organising maps. *Benchmarking*, 11, p. 267-86.
- Lee, K. C., Han, I. and Kwon, Y. (1996). Hybrid neural network models for bankruptcy predictions. *Decision Support Systems*, 18, 63-73.
- Magnusson, C., Arppe, A., Eklund, T. and Back, B. (2005). The language of quarterly reports as an indicator of change in the company's financial status. *Information & Management*, 42, 561-570.
- Neaupane, K.M. and Adhikari, N.R. (2006). Prediction of tunneling-induced ground movement with the multi-layer perceptron. *Tunneling and Underground Space Technology*, 21, 151-159.
- Odom, M.D. and Sharda, R. (1990). A Neural Network Model for Bankruptcy Prediction. *IJCNN International Joint Conference on Neural Networks*, 2, 163-168.
- Olson, D.L., Delen, D. and Meng, Y. (2012). Comparative analysis of data mining methods for bankruptcy prediction. *Decision Support Systems*, 52(2), 464-473.
- Rose, P. S., Andrews, W. T. and Giroux, G. A. (1982). Predicting Business Failure: A Macroeconomic Perspective. *Journal of Accounting, Auditing & Finance*, 6, 20-31.
- Rumelhart, D.E., and McClelland, J.L. (1989). *Parallel distributed processing: exploration in the microstructure of cognition*, (Vol.1). MA:MIT Press.
- Shin, K.S., Lee, T.S. and Kim, H.Y. (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, 28, 127-135.
- Tam, K.Y. and Kiang, M.Y. (1992). Managerial applications of neural networks: The case of bank failure predictions. *Decision Sciences*, 38, 926-948.



Tarassenko, L., Hayton, P., Cerneaz, N. and Brady, M. (1995). Novelty detection for the identification of masses in mammograms. *In: Proceedings fourth IEE International Conference on Artificial Neural Networks, Cambridge*, pp.442-447.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. New York: Springer-Verlag.

