

COMBINATION OF DISCRIMINANT ANALYSIS AND ARTIFICIAL NEURAL NETWORK IN THE ANALYSIS OF CREDIT CARD CUSTOMERS

Mehmet Yazıcı¹

Istanbul Arel University, Turkey

Email: mehmetyazici@arel.edu.tr

ABSTRACT

The decrease in the rate of legal proceedings for the bank loans in the last two years caused an increase in the importance of efficiency along with a decrease in financial resources. When it is paid attention to the distribution in the rate of non-performing loan, it is seen that the most important share belongs to credit cards. Efficient risk management in banking and efficient use of resources depends on taking the necessary actions today by forecasting the potential risks. This study is directed to quick and correct assessment of the credibility of credit card customers. The objective of this study is to present an alternative approach regarding the assessment of credit card customers by using discriminant and artificial neural network methods together. The analysis, applied to the data set which is comprised of 133 samples and being comprised of discriminant and artificial neural network combination, was found to be statistically significant.

Key Words: Credit Cards, Banking, Financial Crisis, Discriminant Analysis, Artificial Neural Networks.

JEL codes: C44, C45, G21.

I. INTRODUCTION

Efficient management of risk in banking and efficient use of resources depends on taking the necessary actions today by forecasting the potential risks.

¹ Mehmet Yazıcı is an Assistant Professor at Istanbul Arel University, School of Applied Sciences, Department of Banking and Finance, Türkoba Mah. Erguvan Sok. No:26/K, Tepekent-Büyükçekmece/İstanbul, Turkey.

According to Central Bank's data of September 2010, personal loans constitute 26.5% of total loans and credit cards have a proportion of approximately 18% among the personal loans (www.tcmb.gov.tr, 20.12.2010). However when we look at the receivables as liquidation, found that the proportion of credit cards is very high and it increased from the value of 53% in March to 57% in September.

26 million people still have 46.689.614 numbers of credit cards in Turkey (<http://www.bkm.com.tr>). According to data of BDDK, 2,282,000 credit card customers are at the process of legal proceeding. Tefvik Bilgin who is the chairman of BDDK attributed the proceeding rate which increased to 10.4% in credit cards to "bankers not paying attention to customer credibility to gain loan volume and consumer fault" in one of his speeches².

This study is directed to quick and correct assessment of the credibility of credit card customers. Many studies have been made the prediction of financial failure. Some of which involve various methods, such as discriminant analysis, logistic regression, decision trees, artificial neural networks and combination of them were research subjects.

In the second part of this study which is comprised of five parts, aim, scope and method are stated. In the third part information were given concerning the applications of discriminant analysis and artificial neural network; in the fourth part an experiment for decision support system being comprised of the combination of discriminant analysis and artificial neural network methods directed to separation of problematic and unproblematic customers among the credit card customers was made. Finally in the last part findings acquired and result were discussed.

II. AIM, SCOPE AND METHOD OF RESEARCH

Prediction of non-performing loans has a great importance in terms of the profitability and productivity in banks. Especially in the area of credit card which necessitates taking quick and correct decision, an increase in legal proceeding rates and reaching to half of total number of proceedings necessitates an improvement in this area. The aim of this study is to develop an alternative method for risk assessment of credit card customers. Accordingly by decreasing the number of used variables, shortening the assessment and decision-taking process and increasing the hit rate in credit cards are aimed with the purpose of accelerating the

² Taken from the speech of Tefvik Bilgin who is the chairman of BDDK made in the conference about "Turkish Banking Sector Current situation and Expectations" in Süleyman Demirel Cultural Centre prepared by Social Sciences Vocational School of Selçuk University and Young Financiers Student Group.

decision-taking process. The applied method is the first example in the assessment of credit card requests.

133 credit card customers selected randomly among the customers of a commercial bank with private capital were analyzed to distinguish between good and bad credit card customers. 11 of 23 independent variables were quantitative and 12 of them were qualitative. 23 of 133 credit card customers were problematic and 110 of them were unproblematic. Customers who do not make payment to credit card for 3 months and over were defined as problematic. Problematic customers were showed as 0 and unproblematic customers were shown as 1 on the data set.

Data set, comprised of 24 variables and 133 samples, was firstly subjected to discriminant analysis and variables obtained after this analysis and were found to be statistically significant and the score calculated according to these variables were subjected to artificial neural network analysis. Data set comprised of 133 samples were randomly divided into two parts, 97 sets as training group and 36 sets as test group³.

III. DISCRIMINANT ANALYSIS AND ARTIFICIAL NEURAL NETWORK

It is a technique which helps to predict the membership appointment with the help of discriminant function which consists of independent variables affecting the determination of group membership mostly. The most famous one of the studies made with discriminant analysis is the study of Altman (Altman, 1968).

Discriminant analysis is a statistical technique used for classifying the units as depending on the feature of (n) number by minimizing the classification error possibility (Hair, Rolph, Tatham, 1998).

Discriminant function;

It is stated as $Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 1.0X_5$

Here;

X_1 = Working capital divided by total assets;

X_2 = Retained earnings divided by total assets

X_3 = Earnings before interest and taxes (EBIT) divided by total assets

X_4 = Market value divided by total debt

X_5 = Sales divided by total assets

Z = Discriminant value.

The process of separation after acquiring the discriminant function is as in the multiple regression. After the parameters of multiple regression

³SPSS Statistics 17.0 package program was used in analyses.

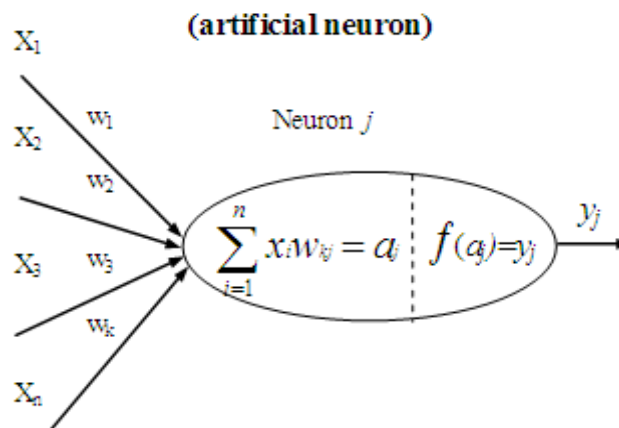
and discriminant analysis are calculated, it is possible for the value of dependent variable to get over the 0-1 interval.

Whatever the value of Z of independent variables is, maintaining it between the 0-1 intervals is possible by using a cumulative probability function. Logit cumulative probability function can solve this function (Maddala, 1988).

Discriminant analysis provides an effective method of discriminating groups from each other by using mathematical techniques. Stepwise method used in the application of discriminant analysis is the method of Wilk's Lambda.

Artificial neural network is information processing system working similar to the characteristics of biologic neural networks. They are computer systems which have been developed to automatically perform some talents without any help such as reproducing new information by means of learning, creating new information and being able to discover that are the characteristics of human brain (Öztemel, 2003).

Figure 1: Artificial Neuron



The ability of artificial neural networks to make parallel operation and features of learning and generalization presents us these algorithms as a very appropriate alternative for credit assessment operations. The first level in artificial neural network approach is to decide on which variables are important in prediction of financial failure. The model presenting the input-output relations between variables are created in this level. After then artificial neural networks are designed with respect to this model and the ability of learning is brought to artificial neural network by being used data in the sample. After providing the ability of learning, predictions are obtained from artificial neural networks.

VI. COMBINATION OF DISCRIMINANT ANALYSIS AND ARTIFICIAL NEURAL NETWORK

Many studies, comparing of artificial neural network and statistical method and the most successful example, were made and the most successful example is the one belongs to Altman, Giancarlo and Franco. In this study, 1000 Italian companies (studied between the years of 1982 and 1992) were subjected to separation analysis with artificial neural network and statistical techniques. A statistical significance was not found between artificial neural networks and statistical methods and it was stated that hybrid systems using combining both methods may give better results (Altman, Giancarlo, Franco, 1994).

Some combination experiments were made in the literature within this context and successful examples were presented by researchers such as Widder, Ammon, Schaeffer and Wolff (2008) and Chen, Xun, Li and Zhang (2010). Discriminant analysis and artificial neural network implementation were applied to randomly selected 133 credit card customers of a bank in parallel with these studies.

Table 1. Eigen Values

Function	Eigen Value	% of Variance	Cumulative%	Canonical Correlation
1	1,443	100,0	100,0	,768

It look at Canonical Correlation, Eigen Value and Wilk's Lambda statistics to determine how much important the discriminant function is. Canonical Correlation measures the relation between discriminant scores and groups and shows the explained total variance. Canonical correlation coefficient in Table 1 was found 0,768. We should square to decipher this value $(0,776)^2 = 0,59$. In other words our model can explain 59% of variance in the dependent variable.

The bigger the Eigen Value statistic is, the more of variance in the dependent variable can be explained by that function. As well as it is not a definite value, Eigen Values bigger than 0,40 are accepted as good. The Eigen Value in the study was found 1,443. This value is an acceptable value. Because dependent variable has two categories, there will only be one discriminant function.

Table 2. Wilks' Lambda

Tested Function	Wilks' Lambda	Chi-square	Df (Degree of freedom)	Sig.(Significance probability)
1	,409	113,415	8	,000

Wilks' Lambda statistic in Table 2 gives the part of total variance in discrimination scores not explained by differences between groups. In the study, approximately 40% of total variance in separation scores is explained by differences between groups.

Table 3. Standardized Canonical Discriminant Function Coefficients

	Function
	1
Income/ Expense (%)	,284
Occupation	-,303
Marital Status	-,333
Security Possession	,328
KKB Score	,706
Unpaid Cheque /Dishonored Promissory Note Inquiry	,494
Legal proceeding (Last 3 years)	,285
Irregular payment	,375

As can be seen in Table 3, 9 of 23 independent variables were found to be statistically significant and included in function.

Table 4. Separation Results

Situation of Credit Card	Prediction of Group Membership		Total
	0	1	
Original Number 0	20	3	23
1	9	101	110
% 0	82,6	17,4	100,0
1	8,2	91,8	100,0

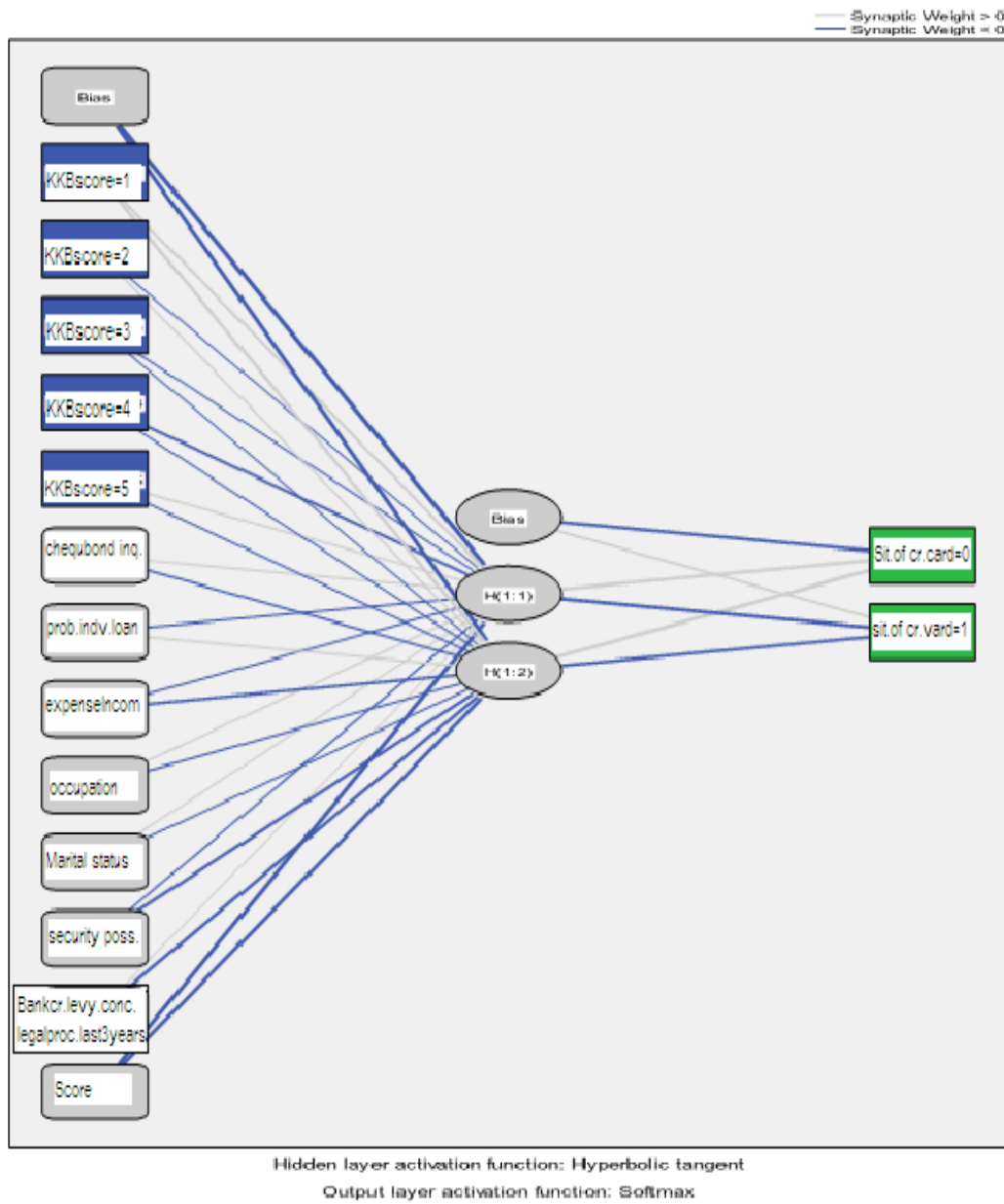
According to separation analysis, group membership occurred in the ratio of 87.2% rightly on average. Scores, calculated by being used the data of 9 dependent variables which was found to be significant in appointment of group membership and the obtained variable weights after the separation analysis, were used as input variables in the analysis of artificial neural network. Usage of statistically significant variables as input in the analysis of artificial neural network ensures being predictive in credit decisions.

Table 5. Data Set of Artificial Neural Network

		Number	%
Sample	Training	97	72,9%
	Test	36	27,1%
Valid		133	100,0%
Invalid		0	
Total		133	

133 credit card customers who were included in the analysis were randomly divided into Training and Test groups by SPSS statistics program. After this discrimination, it was determined that there was 97 sample training group and 36 sample test group.

Figure 2. Feed-forward Artificial Neural Network



Following the analysis, group memberships occurred 100% correct in the training group and 91.7% in test group.

Table 6. Separation

Sample	Observation	Prediction		
		0	1	True Separation%
Training	0	14	0	100,0%
	1	0	83	100,0%
	Total %	14,4%	85,6%	100,0%
Test	0	6	2	75,0%
	1	1	27	96,4%
	Total %	19,4%	80,6%	91,7%

IV. FINDINGS AND CONCLUSION

Usage of a large number of variables in credibility analyses causes the loss of capital and time in terms of collection and storage of this information. Besides many data that not significant in statistical terms and do not contributing to prediction power of model are included in the analysis. When we evaluate in terms of bank allocation, we find that while the usage in the management of customer relations is significant, these information create a limited effect after a certain variable during taking credit decisions. While a decrease in the number of significant variables which will be used with discriminant analysis ensures time and capital savings, it increases the incidence percentage in decision in case it is used as input in methods of analysis such as artificial neural network.

The aim of this study is to prevent the loss of profit and efficiency of banks caused by high legal proceeding numbers in credit cards by minimizing the loss of efficiency experienced in allocation of credit cards. For the study is tested on data of limited number of customers, increasing this number in future studies will ensure that the results to be acquired will be more healthier. The results obtained from this study, which is a first experiment made in this field, are promising for future.

Decision support models which are comprised of statistical methods and combination of advanced methods such as artificial neural network, fuzzy logic and genetic algorithm can be used efficiently in some fields such as credit card allocation in which quick decision making is necessary.

REFERENCES

- ALTMAN E.I. (1968), Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *The Journal of Finance*, V: 23 n: 4 pp. 589-609.
- AKTAŞ R., DOĞANAY M. M., YILDIZ B. (2003), Prediction of Financial Failure: Comparison of Statistical Methods and Artificial Neural Network, *Ankara University Faculty of Social Sciences Journal*, V:58, Number: 4 p.6.
- CHEN Xiao, XUN Yi, LI Wei and ZHANG Junxiong (2010), Combining discriminant analysis and neural networks for corn variety identification, *Computers and Electronics in Agriculture*, Volume 71 pp. 48-53.
- COAST K.P., FANT F.L. (1993), Recognizing Financial Distress Patterns Using a Neural Network Toll, *Financial Management*, pp. 142-155.
- HAIR J., ROLPH A.E., TATHAM W.C. (1998), *Multivariate Data Analysis*, Printice-Hall International, pp. 239-326.
- KESKİN Y. (2002), Prediction of Financial Failure in Companies, Multiple Model Suggestion and Application, *Hacettepe University Institute of Social Sciences Department of Business Dissertation*, pp. 43-114.
- MADDALA G.S. (1988), *Introduction of Econometrics*, Newyork, McMillan Publishing Company, pp.16-32.
- ÖZTEMEL E. (2003), *Artificial Neural Networks*, Papatya Publishing.
- WIDDER A., AMMON R.V., SCHAEFFER P. and WOLFF C. (2008), Combining Discriminant Analysis and Neural Networks For Fraud Detection on the Base of Complex Event Processing, *2nd International Conference on Distributed Event-Based Systems*.
- WILKSON R.L., SHARDA R. (1994), Bankruptcy Prediction Using Neural Networks, *Decision Support Systems*, V:11 pp. 545-557.
- YILDIZ B. (1999), Usage of Artificial Neural Network in Prediction of Financial Failure and an Empirical Study, *Dumlupınar University Institute of Social Sciences Department of Business Dissertation*, pp. 14-155.
- http://www.bkm.com.tr/istatistik/pos_atm_kart_sayisi.asp, 21.06.2010.
- <http://www.cnnturk.com/2010/ekonomi/genel/03/31/2.milyon.282.bin.kredi.karti.musterisi.takipte/570323.0/index.html>, 20.05.2010.
- <http://debs08.dis.uniroma1.it>, 10.06.2010.
- <http://www.bddk.org.tr>, 20.10.2010.
- <http://www.tcmb.gov.tr>, 20.12.2010.
- <http://www.cnnturk.com>, 22.11.2010.