INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE STUDIES
Vol 6, No 1, 2014 ISSN: 1309-8055 (Online)

ON THE PERFORMANCE OF ARTIFICIAL INTELLIGENCE METHODS FOR FAILURE PREDICTION: EVIDENCE FROM ISTANBUL STOCK EXCHANGE

Dr Hakan Er
Akdeniz University, Department of Business Administration
Dumlupinar Bulvari, Kampus, Antalya 07058 TURKEY
Phone: +90 242 310 22 26
E-mail: erhakan@akdeniz.edu.tr

─Abstract ─

The prediction of business failure is a widely studied subject in financial literature. Many earlier studies on this topic employed statistical methods such as multiple discriminant analysis, logit and probit to predict corporate failure using past financial data (especially the ratio data). However, there has been a recent surge in academic interest in the use of artificial intelligence (AI) methods to predict financial distress. Numerous studies documented that AI methods outperform traditional methods. Majority of these studies used data from established markets, the number of studies on emerging market data is rather limited and only a handful of studies employed Turkish data for analysis. This study aims to contribute to the literature by applying the artificial neural networks to predict deletions from Istanbul Stock Exchange (ISE) National 100 Index. The sample is constructed using the quarterly fundamental data of the companies listed in this index the period between January 2008 and December 2012. We employed Neural Networks (NN), logit and probit to predict deletions from index one quarter before they have occurred. Results show that although the logit provides slightly better in-sample predictions, all of the methods fail to identify deletions in the out-of-sample periods.

Key Words: *Neural Networks, Genetic Programming, Business Failure Prediction, Emerging Markets*
JEL Classification: C45, G33

1. INTRODUCTION

Business failure prediction has long been an important and widely studied topic. The emergence of academic interest on corporate failure prediction date back to late 60's. In his seminal work, Altman (1968) used financial ratios to develop a "Z-Score" which forecasts the probability of a company entering bankruptcy within a two year period. This study provided the basis for the methodology of numerous other studies that followed. Early studies were in the USA, later on studies in other developed countries such as United Kingdom, Germany and Japan appeared. There have also been many studies from authors from emerging or developing economies. Early studies used financial ratios to build early warning systems for business failure based on statistical models. Methods such as multiple discriminant analysis, logit and probit have been the most widely employed statistical methods for failure prediction and later on became benchmarks for the failure prediction problem. In the last two decades we have witnessed the growing use of the artificial intelligence methods for the development of failure prediction models in the financial literature. Due to space limitations we are unable to provide an extensive review of the studies on this field. Altman and Narayanan (1997) present a comprehensive coverage of the earlier corporate failure research and Kumar and Ravi (2007) present a more up to date review focusing more on studies involving artificial intelligence methods.

In Altman and Narayanan's (1997) review failure prediction studies have been classified into two subgroups: studies in developed and in developing countries. Authors provide five main characteristics of developed country models:

- failure prediction studies have a long history,
- corporate financial data are more readily available,
- failure is easier to identify because of the existence of bankruptcy laws and banking infrastructures
- government intervention is somewhat less, but not nonexistent
- there is a more sophisticated regulation of companies to protect investors.

Developing country models lack at least one of these characteristics. Turkey, the focus of this study, should be considered as a developing country according to these criteria (at least according to the first of them). Despite the large number of failure prediction studies conducted in many other countries, we have only identified only a handful of studies using Turkish data. This provides the first motivation for this paper. This paper aims to contribute to the growing literature on failure prediction by providing evidence from this important emerging market.

As mentioned earlier, a common feature of failure prediction studies covered is the use of financial ratios for model building. However, they differed in the way in which the failure is described. For the majority of the studies, the failure meant bankruptcy. However, bond default, bank loan default, delisting from a stock exchange and even deletion from an index were considered as failure by many other studies. The last failure definition is the one that has attracted the least attention in the literature.

A company's addition to or deletion from an index can have a significant impact on its return and its liquidity. After the announcement of an index reconstitution, the value of the stocks added to (deleted from) tend to increase (decrease) resulting positive (negative) abnormal returns (Lynch and Mendelhal, 1997, Meera et al, 2000, Mase, 2007). Addition (deletion) also brings positive (negative) liquidity effects. Moreover, each time a firm is added to and deleted from an index, index composition changes. This also presents important cost implications for the portfolio managers who track the index. Therefore, building a model for the prediction of the deletions from an index is an important task. Although there are many studies on the consequences of deletions on both the stock deleted and on the index it deleted from, the number of the studies on the models for the prediction of stocks' deletion are rather limited. This provides the second motivation for this paper. This study will also contribute to failure prediction literature by providing evidence on the prediction models developed for index addition/deletion.

The rest of the paper is organized as follows: Section two provides brief information on logit and probit models. Section three summarizes neural networks. The details on the data and the methodology utilized in this study are provided in Section three. Section four summarizes the results. Conclusion and suggestions for future work are presented in final section.

2. LOGIT AND PROBIT MODELS

The logit/probit models are the most commonly applied approaches when the response variable in a regression is dichotomous. As the dependent variable in failure prediction studies is dichotomous and as these techniques are easy to interpret they are widely used in academic studies on failure prediction. One other reason for their popularity is that these models avoid many of the statistical problems of alternative models that impose statistical requirements on the distributional properties of the predictors.

In failure prediction studies, both the logit and probit techniques are used to create a score (a probability) for each firm by weighting the independent variables. Based on that score a firm then is classified as failed or non-failed, using a cut-off probability.

In the logit model the probability of failure is given by the following cumulative logistic function:

$$P_i = \frac{1}{1 + e^{-Z_i}} = \frac{e^{Z_i}}{1 + e^{Z_i}} \qquad \textbf{(1)}$$

Where $Z_i = \beta_1 + \beta_2 X_{1i} + \beta_3 X_{2i} + \cdots + \beta_n X_{ni}$

The logit model is a very similar to single neuron network (neural networks will be explained in the following section).

The only difference between the logit and the probit models is the way the probability of failure is calculated. If the same Z score is placed into a cumulative normal probability function the model is called a probit model.

$$P_i = \int_{-\infty}^{Z_i} \frac{1}{2\pi} e^{-Z/2} dx \qquad \textbf{(2)}$$

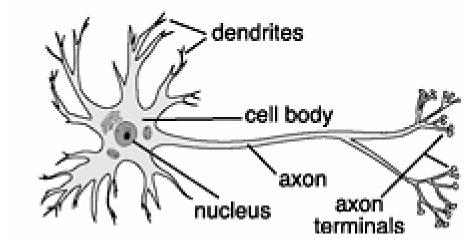
The functions for P_i's are also called the link functions. Logistic has slightly flatter tails. Probit curve, on the other hand, approaches the axes more quickly than the logit curve. There is no generally accepted rule to select probit or logit. When one is concerned with tail part of the curve, the selection of logit or probit might matter. In this study both of these models will be utilized.

3. ARTIFICIAL NEURAL NETWORKS

Many different artificial intelligence approaches such as inductive learning, NNs, GAs and case-based reasoning (CBR) have been utilized in model building for failure prediction. However, the research on failure prediction utilizing NNs started at the beginning of nineties and currently they are the most commonly employed methodologies for classification problems in failure prediction. There are also many commercial loan default prediction products which are based on NNs (Atiya, 2001).

The biological findings relating to the behavior of human brain as a network of units called "neurons" (McCulloch and Pitts, 1943) led to the emergence of mathematical models called Artificial Neural Networks in the early 60's. Neurons are the basic signaling units of human nervous system. Each neuron is a discrete nerve cell and consists of a cell body (or soma), an axon and the dendrites. The cell body contains the nucleus and maintains the protein synthesis. Dendrites receive signals from other neurons.

Figure 1:

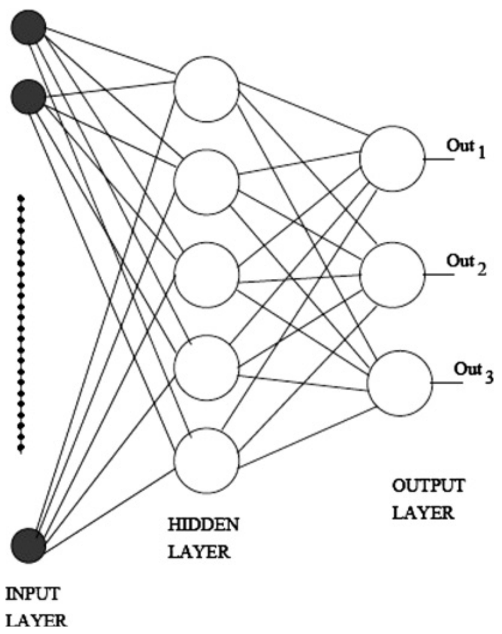


Source: http://science.education.nih.gov/supplements/nih4/self/guide/info-brain.htm

There are around 100 billion neurons in the human brain and each neuron is connected to thousands of other neurons through synapses. The synapse is the area of contact between two neurons. Patterns of chemical and electrical signals travel between the neurons. When a neuron is activated an electrical current is transmitted down its axon and reaches to an axon terminal. Here is where the neurotransmitters (brain chemicals) is released. The neurotransmitter eventually pass through synapse and reach to a receptor in the dendrite of another neuron which in turn passes transmitter to another neuron. When the same connections are reactivated so many times, the synapse changes physically. The connections become more and more efficient. This is briefly how the experience or behavior is stored in our long-term memory.

Artificial neural networks are a broad class of models that mimic functioning of the human brain explained above. An NN is a set of interconnected input, processing and output units called neurons. NNs can be classified into two types: recurrent and feed-forward NNs. In recurrent NNs multiple neurons can be interconnected. Feed-forward NNs, on the other hand, is composed of an input layer; one or more hidden layers and an output layer. .

Figure 2: A One Layer Feed-forward NN



In feed-forward backpropagation networks, which are are the most commonly employed in financial research, neurons are not connected to each other within a layer. A neuron in the input layer is connected to neurons only in the next hidden layer. If there is only one hidden layer between the input and output layers, the results from each neuron in the input layer is first send to the neurons in the hidden layer. Then each neuron in the hidden layer sends its results to the neuron/neurons in the output layer. If the network consists of more than one hidden layers, each neuron in the layer immediately after the input layer receives the results from the first layer and then sends its results to next hidden layer.

A feed forward NN is characterized by the number of input neurons, the number of hidden layers, the number of neurons in each hidden layer, the number of output neurons and weights for all connections. The number of input neurons and the number of output neurons are decided by user and therefore are exogenous to the model and there is no fixed rule to decide on the number of hidden layers and the number of neurons in hidden layers. Trial and error usually dictates the decision.

Prediction for *n* observations starts with a random set of weights. The first observation is fed forward through the network using these weights and the transfer function. The "weights" transform the input and send the resulting data to the neurons in the next layer via the following function:

$$O_j = G\left(\sum_{i=1}^{m} w_{ij}x_{ij}\right) - \theta_j \qquad \textbf{(3)}$$

Here O_j is the value of the jth neuron in the layer, x_{ij} is the ith neuron in the previous layer, w_{ij} is the weight on the connection from ith neuron in the previous layer, m is the number of neurons in the previous layer and θ_j is the bias of the neuron. The function *G( )* is called the activation (or the transfer) function. Although, there are various choices for this function, a sigmoid hyperbolic tangent function $G(z) = \tanh(z) = \frac{1-e^{-z}}{1+\mathrm{e}^{-z}}$ is a commonly used activation function for applications in backpropagation networks.

The prediction of the model, Y_1, is then compared with the actual output (desired output or target), D_1, to compute the error, E_1. The weights are then adjusted to reduce the error for the first observation. This procedure is repeated for the remaining observations. One cycle is completed when the final observation is reached. Many such training cycles are repeated until the weight vector, $\boldsymbol{W}$, which minimizes the overall prediction error, E, is obtained.

In back-propagation learning algorithm weight adjustment process starts at the output units and works back to the first hidden layer. Overall error $\sum_{i=1}^{n}(D_i - Y_i)^2$ is a function $\boldsymbol{W}$:

$$E(W) = \sum_{i=1}^{n}(D_i - Y_i(W))^2 \qquad \textbf{(4)}$$

Each individual weight update, Δw, is then calculated according to following formula:

$$\Delta w = \propto \frac{\partial E}{\partial w} \quad \textbf{(5)}$$

The new weight is the old one plus the update. α (between 0 and 1) is a gain term which is also called the learning rate. The learning rates are parameters in the learning rule that aid the convergence of error. Adding a momentum term (again between 0 and 1) and smoothing weight changes may sometimes result in a faster convergence.

Ideally this procedure stops when we reach the global minima of the error surface. However, an optimal solution is not guaranteed and there is no way of knowing whether the global minima have been found or not. The procedure is usually stopped when the decrease in the total prediction error is small or the overall changes in the weights are negligible.

4. DATA, METHODOLOGY AND RESULTS

The sample used for empirical analysis is constructed from the data on companies listed in Istanbul Stock Exchange 100 (ISE 100, now Borsa Istanbul 100 Index, BIST 100) Index in the period between January 2008 and December 2012. Banks and other finance firms, insurance companies, holding companies and investment companies are excluded.

All of the firms added to and deleted from the index within this period are identified. The reasons for the deletions are also investigated. Deletions as a result of the mergers and acquisitions, takeovers are ignored.

The sample is constructed as follows: All of the firms listed in ISE 100 in the first quarter of 2008 are identified. The ones that are deleted from index in the second quarter of 2008 are considered as the unsuccessful firms (deletions) and coded as "1". The ones that are still listed in the index are considered as successful and their dependent variables are coded as "0". The fundamentals of all firms listed in the index in the first quarter are collected and the financial ratios used in the models are calculated. The procedure is repeated until the last quarter of 2012. The fundamentals of the stocks in the sample are collected from www.isyatirim.com. This procedure yielded 812 observations and 27 of these observations are identified as deletions.

The sample is then divided into training and validation sets. All observations from the period between the first quarter of the 2008 and the last quarter 2012 are included in the training set. The remaining observations are included in the validation set. The training set yielded 470 observations and 14 of these observations are on deleted companies. The validation set, on the other hand, has 342 observations and 13 of them are on deleted companies.

The following financial ratios are calculated and used as dependent variables in logit and probit models and as inputs in the NN:

- (Cash and Cash Equivalents + Marketable Securities) / Current Liabilities
- Cash and Cash Equivalents / Current Liabilities
- Asit Test Ratio: (Current Assets – Inventory) / Current Liabilities
- Current Ratio : Current Assets / Current Liabilities
- Current Assets / Total Assets
- Fixed Assets / Total Assets
- Net Working Capital / Total Assets
- Debt Ratio: Total Liabilities / Total Assets
- Long Term Liabilities / Total Assets
- Net Working Capital Turnover: Net Sales / Net Working Capital
- Credit Turnover: Net Credit Sales / Average Accounts Receivable
- Accounts Payable Turnover: Cost of Sales / Average Accounts Payable
- Net Sales / Long Term Liabilities
- Net Income / Total Assets
- Net Income / Shareholder Equity
- Log of Total Assets

The results for the training set of the logit and probit and logit models using these ratios as predictors are shown in Table 1.

Table 1: Logit and Probit Results						
	LOGIT			**PROBIT**		
Variable	**Coefficient**	**Standard Error**	**Z**	**Coefficient**	**Standard Error**	**Z**
const	4.85E+06	NA	NA	4.17E+06	1.84E+07	0.2269
Current Ratio	1.3767	0.883898	1.558	0.588512	0.360431	1.633
Quick Test Ratio	0.397533	0.973751	0.4082	0.124361	0.46557	0.2671
Cash Ratio	0.0845063	1.08419	0.07794	0.0102532	0.434646	0.02359
NWC/Total Assets	4.85E+06	NA	NA	4.17E+06	1.84E+07	0.2268
ROA	6.38144	6.43931	0.991	2.8038	2.81816	0.9949
Fixed Assets/Total Assets	4.85E+06	NA	NA	4.17E+06	1.84E+07	0.2268
ROE	0.149311	0.111433	1.34	0.0595128	0.0484358	1.229
Debt Ratio	4.85E+06	NA	NA	4.17E+06	1.84E+07	0.2268
NWC Turnover	0.000889368	0.00089	0.9993	0.000401884	0.0004217	0.9531
Long Term Liabilities/Total Assets	4.85E+06	NA	NA	4.17E+06	1.84E+07	0.2269
Credit Turnover	0.0498551	0.0468139	1.065	0.0231344	0.0208984	1.107
Payables Turnover	0.0316459	0.0213452	1.483	0.0146432	0.0113594	1.289
Log Total Assets	0.682774	0.23229	2.939	0.312368	0.104729	2.983

Due to perfect colliniearity three predictors omitted. Apart from the “Log of Total Assets” none of the coefficients are signicant in either models.

The logit and probit’s recommendations for test period are calculated using these coefficients. An NN is also run and its contingency table along with the contingency tables of logit and probit are shown in Table 2:

Table 2: Contingency Tables						
		Training Set			Validation Set	
	TRUE	Predicted		TRUE	Predicted	
		1	0		1	0
NN	1	0	14	1	0	13
	0	0	470	0	0	315
Rate of Correctness			97.11%			99.06%
LOGIT	TRUE	Predicted		TRUE	Predicted	
		1	0		1	0
	1	1	13	1	0	13
	0	0	470	0	0	315
Rate of Correctness			97.31%			99.06%
PROBIT	TRUE	Predicted		TRUE	Predicted	
		1	0		1	0
	1	0	14	1	0	13
	0	0	470	0	0	315
Rate of Correctness			97.11%			99.06%

5. CONCLUSION

The results of this study show that prediction of deletion from an index is quite a challenging task. There are only 27 deletions in the sample and even if the sample period is expanded the ratio of the deletions will not change dramatically. Although, the methods used in this study are among the most popular methods in failure prediction and it utilizes the most commonly employed predictors in failure prediction models, none of the methods can identify any deletions in the training period.

In future work predictors specific to this prediction problem will be identified other statistical methodologies like survival analysis and alternative artificial intelligence methods like genetic algorithms will be utilized.

BIBLIOGRAPHY

Altman, E. I. (1968), Financial Ratios, Discriminant Analysis And The Prediction Of Corporate Bankruptcy, *Journal of Finance,* Vo. 23, No. 4, pp. 589-609.

Altman, Edward I., and Paul Narayanan, (1997), An international survey of business failure classification models, *Financial Markets, Institutions & Instruments,* Vol. 6, No. 2, pp. 1-57.

Atiya, Amir F. (2001), Bankruptcy prediction for credit risk using neural networks: A survey and new results, *Neural Networks, IEEE Transactions on* Vol. 12, No: 4, pp. 929-935.

Kumar, P. R., & Ravi, V. (2007), Bankruptcy prediction in banks and firms via statistical and intelligent techniques – A review, *European Journal of Operational Research*, Vol. 180, No. 1, pp. 1–28.

Lynch, A. W., and R. R. Mendenhall. (1997), New Evidence on Stock Price Effects Associated with Changes in the S&P 500 Index, *Journal of Business* Vol. 70, No. 3, pp. 351-383.

Mase, B. (2007), The Impact of Changes in the FTSE 100 Index, *The Financial Review* Vol. 42, No. 3, pp. 461-484.

Meera, Ahamed Kameel, Niranjan Tripathy, and Michael R. Redfearn (2000), Wealth and liquidity effects of stock delistings: empirical evidence from the stock exchanges of Singapore and Malaysia, *Applied Financial Economics*, Vol. 10, No. 2, pp. 199-206.