FUNDAMENTAL ANALYSIS WITH ARTIFICIAL NEURAL NETWORK

Birol Yildiz, Eskisehir Osmangazi University
Ari Yezegel, Bentley University

ABSTRACT

This study performs fundamental analysis and cross-sectional prediction of stock return with neural network technology. Eighteen financial ratios are used as the input vector and one-year ahead stock returns are used as the output vector. The fundamental analysis trading strategy generated by artificial neural networks yields an average annual abnormal return of 22.32% after controlling for market risk, book-to-market, size and momentum effects. Our results highlight neural network's ability to predict future returns in NYSE/AMEX/Nasdaq securities for the period 1990-2005. Artificial neural network technology stands out as a valuable tool for fundamental analysis and forecasting equity returns in the U.S. markets.

JEL: G11; M41; C45

KEYWORDS: Fundamental Analysis, Stock Market, Neural Network

INTRODUCTION

Fundamental analysis is an irreplaceable part of financial analysts', institutional and individual investors' toolbox (Carter and Van Auken 1990). Consistent with the high demand for fundamental analysis, prior literature documents evidence emphasizing fundamental analysis' predictive power of future returns (Ou and Penman, 1989; Lev and Thiagarajan, 1993; Abarbanell and Bushee, 1997, 1998; Greig, 1992; Houltsen and Larcker, 1992). Prior literature testing fundamental analysis's ability to generate future abnormal return assumes the relationship between financial ratios and returns to be linear. However, there is strong evidence indicating the relationship between key fundamental ratios and returns to be non-linear (Beaver, Clerke and Wright, 1979; Freeman and Tse, 1989; Abdel-khalik, 1990; Pahor and Mramor, 2001; Omran and Ragab, 2004). Nevertheless, prior studies rely on a linear estimation technique to measure fundamental analysis's predictive power of returns. Relying on a linear estimation method restrains the relationship to a linear one which limits the predictive power of fundamental analysis.

We examine whether a statistical method that does not assume a linear relationship generates significant returns adjusted for the beta and other previously documented anomalies such as the size, book-to-market and momentum anomalies. To test whether a non-linear assumption generates significant returns we propose the use of neural network technology.

The most powerful feature of artificial neural network technology is solving nonlinear problems that other classical techniques do not deal with. The artificial neural network (ANN) technology does not require any assumption about data distribution and missing, noisy and inconsistent data do not possess any problems. Another important aspect of the ANN is its ability to learn from the data. Artificial neural network technology is used in classification, clustering, predicting, forecasting, pattern recognition problems successfully. In most cases, ANN technology produces superior performance than other statistical techniques (Haykin, 1999; Hill, Marques, O'Connor and Remus, 1994; Cheng and Titterington, 1994; Zhang, Patuwo and Hu, 1998).

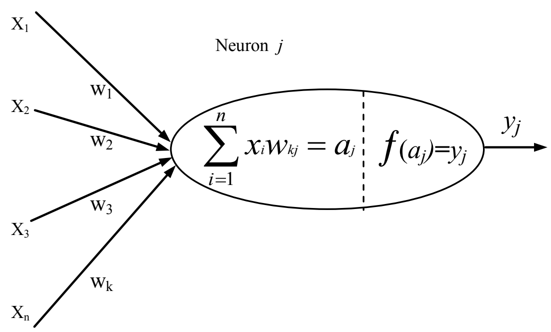
ANN technology has been commonly used in accounting and finance literatures specifically in bankruptcy prediction, bond rating, stock index forecasting, foreign exchange rate forecasting, credit evaluation, fraud detection (Wong, Bodnovich and Selvi, 1997; Wong and Selvi, 1998; Vellido, Lisboa and Vaughan, 1999; Krishnaswamy, Gilbert, and Pasley, 2000; Wong, Lai and Lam, 2000; Coakley and Brown, 2000). However, there is limited examination of the performance of ANN using U.S. equity data.

The remainder of the paper is organized as follows. The next section introduces the artificial neural network technology. The subsequent section describes the sample and performance measurement methodology. Section 4 presents our empirical results and section 5 concludes.

LITERATURE REVIEW

Artificial neural network is an information processing system, which simulates some human brain functionality like thinking and learning. The first commercial artificial neural network designer was Robert Hecht-Neilsen. He defined artificial neural network as a distributed information processing structure which consisted of simple and interconnected processing elements that produces dynamic outputs for each input (1990:2).

Figure 1: An Artificial Neuron



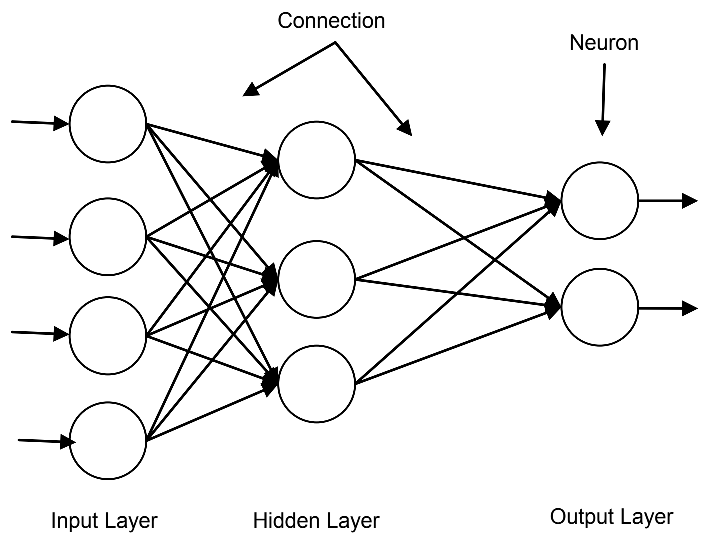
An artificial neuron sums all inputs and produces an output.

The major components of neural network are neurons, connections, and the learning algorithm. A neuron (*j*) is basic processing unit of a neural network. All neurons in the network receive a number of inputs (x_i) and generate an output (y_i). These outputs can be either an input to other neurons or output to outside of neural network (See Figure 1).

A neural network is built by interconnection of neurons (See Figure 2). The term “layer” is used to indicate the row of neurons in the artificial neural network. The first layer is identified as an input layer that receives data from outside of neural network. Output layer is the last layer in the neural network and sends calculated results to the outside. Layers between input and output are called hidden layers.

Data is transmitted among neurons through these connections. Any connection between neuron (*k*) and neuron (*j*) has a weight (*wkj*) and each input is multiplied by its respective weighting factor. This operation is especially important because every input is weighted.

Figure 2: An Artificial Neural Network



An artificial neural network consists of neurons and connections in a particular structure.

Connection types like feed-forward or feedback, number of layers and number of neurons at a layer are identified as architecture. In information processing, data enter input layer and flow on connections as well as neurons through network. In this information processing, data are processed at each neuron. Neurons contain two basic functions to process information: summation function and transfer function (See Figure 1). The summation function (1) gets the weighted sum of all inputs that reach neuron. This function determines stimulation level of neuron.

$$\sum_{i=1}^{n} x_i w_{kj} = a_j \tag{1}$$

The transfer function determines activation level of the neuron and relationship between stimulation level and output (*yi*). The crucial feature of transfer function is limitation of the output.

$$f(a_j)=y_j \tag{2}$$

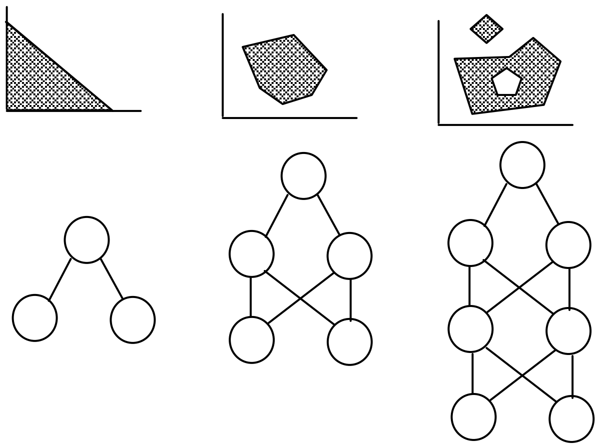
The most popular transfer function is sigmoid functions (3) that limit the output value to between 0 –1 for every input value.

$$y_j = \frac{1}{1+e^{-yi}} \tag{3}$$

Basically, a neural network learns from errors. The learning algorithm calculates errors from the difference between the network output and the desired output that is gathered from real world model. The learning algorithm uses the error term to adjust weights and repeats this procedure until the network produces desired outputs. When the neural network produces desired outputs for all input values it captures real world model that exists between inputs and outputs. Once this phase is completed the neural network is considered trained.

A neural network exhibits some statistical abilities depending on its architecture, especially, when faced with a classification problem whether an input should be classified class A or B. A neural network having threshold transfer function and one connection level can be used to separate the decision space in two categories with a line (See Figure 3.a). A neural network with two connection levels can separate the input space into open convex or close concave planes (See Figure 3.b). If neural network have tree connection levels, it has ability to separate the input space into a number of open or closed planes (See Figure 3.c) (Bishop, 1997:122-124).

Figure 3: Number of Hidden Layer in the Neural Networks and Their Statistical Abilities.



(a) (b) (c)

Neural networks can present some statistical capability according to their structures. (a) The neural network consists only input and output layer can separate input space into two linear pieces. (b) The neural network consist of a hidden layer can separate input space to closed concave and convex piece. (c) The neural network consist of more than one hidden layers can separate input space in to many open and closed pieces. Source- Bishop Christopher M., Neural Networks for Pattern Recognition, Clerendon Press, Oxford, 1997, p.123.

The artificial neural network method has many advantages compared to other techniques (Trippi and Turban, 1996; Schalkof, 1997; Goonatilake and Treleaven, 1995): i) Generalization: The most important advantage of neural network is learning. A trained neural network can reach satisfactory results with incomplete and faulty inputs. For instance, a neural network that is trained to recognize human faces can recognize individuals using photographs that were taken in various conditions (e.g. dark place, a different point of view and etc.). ii) Fault tolerance: Traditional computing systems are very sensitive to faults in systems. Any problem in these systems may cause the system to halt or create an important error in results. However, a neural network is not affected as much as a traditional computing system if some of neurons are damaged. iii) Adaptation: Neural networks can learn and adapt to different environments without requiring the completion of retraining. iv) Parallel distributed processing: All processing units in neural network run simultaneously, so the neural network is speedy. v) No assumption is needed: Artificial neural network makes no assumption about the used data. Any kind of data could be used as input for neural networks. This is the most important advantage of neural network technology.

The artificial neural network method also possesses disadvantages: i) Failure to achieve accurate results: This technology may produce unreasonable and irrelevant results. Sometimes neural networks cannot be trained. ii) Lack of explanation: While other statistical techniques generate understandable and interpretable parameters for problems, neural networks' weights cannot be interpretable. In other words the model used by neural networks remains as a black box. Neural network's features have attracted some interest from researchers and researchers have used neural network analysis to predict future returns and classify stocks to portfolios (Wong, Wong, Goh and Quek, 1992; Kryzanowski, Galler and Wright, 1993).

Nevertheless, prior studies are generally limited in nature: Quah and Srinivasan (1999) applied neural network analysis to Singapore Stock Exchange and Albanis and Batchelor (2007) applied this analysis to a sub-sample of firms listed in the London Stock Exchange.

DATA AND METHODOLOGY

Our sample includes all firms traded in the New York (NYSE), American (AMEX) and Nasdaq exchanges that have data available in both Center for Research in Security Prices (CRSP) and Compustat files. The sample spans over the fiscal years between 1962 and 2005. Companies with missing data are excluded. The final sample contains 136,924 firm-year observations**.** Eighteen financial ratios are calculated for each firm-year (see Table 1). For each firm year we compute the annual buy-hold return beginning four months after the fiscal-year-end and ending twelve months later. We then rank each firm-year into ten groups based on the annual buy-hold return.

Table 1: Financial Ratios

Panel A: Liquidity Ratios		
Current Ratio	Current Assets / Current Liabilities	data #4 / data #5
Quick Ratio	(Cash and Short-Term Investments Plus Receivables – Total) / Current Liabilities	(data #1 + data #2) / data #5
Short Term Debt to Equity	Current Liabilities / Stockholders' Equity	data #5 / data #216
Panel B: Solvency Ratios		
Debt to Equity	Total Liabilities / Stockholders' Equity	data #181 / data #216
Debt to Assets	Total Liabilities / Total Assets	data #181 / data #6
Interest Coverage	Income Before Interest and Tax Expense / Interest Expense	(data #15 + data #170) / data #15
Panel C: Profitability Ratios		
Liquid Asset Turnover	Net Sales / Liquid Assets	data #12 / (data #1 + data #2)
Current Asset Turnover	Net Sales / Current Assets	data #12 / data #4
Tangible Fixed Asset Turnover	Net Sales / Tangible Fixed Asset	data #12 / (data #6- data #4)
Equity Turnover	Net Sales / Stockholders' Equity	data #12 / data #216
Asset Turnover	Net Sales / Total Assets	data #12 / data #6
Gross Profit Margin	Gross Profit / Net Sales	(data #12 – data #41) / data #12
Operating Profit Margin	Operating Profit / Net Sales	data #13 / data #12
Net Profit Margin	Net Profit / Net Sales	data #172 / data #12
Return on Equity	Net Profit / Common Equity	data #172 / data #60
P/E	Price/Earnings	abs(prc)*shrout / data #172
B/M	Book/Market	data #60 / abs(prc)*shrout
P/S	Price/Sales	abs(prc)*shrout / data #12

This table lists the names of the financial ratios used in this study. For each ratio the table reports the formula and the data items that were used to compute the ratio. Variables listed in the form of data #x are from Compustat and prc and shrout variables are from CRSP.

We use financial ratios as the input vector and the twelve months return of the firm as the output vector for developing neural network model. There is no information about which financial ratio is important and we didn't utilize a statistical tool for dimension reduction. We divided our sample into three sub-samples: training, test and validation. Training and testing samples were used to develop the neural network model and assess its performance. For training data, 4,000 random observations were selected from the 1962-1989 sample period. For testing data, we used 70,252 observations from the sample period 1962- 1989. Finally for validation we relied on the sample between the years 1990 and 2005 (66,672 observations).

The most important aspect of the training process is the testing of the neural-network model. In the training process, we used the testing sub-sample to gauge whether there was serious over fitting. In cases where over fitting was a concern we repeated the training process. We finally arrived at a neural network model that is feed forward neural network and has three hidden layer. Each hidden layer consisted of 100, 35, 15 nodes consecutively (Main absolute error, sigmoid (-/+) transfer function and Jacob's Enhanced Back Propagation learning algorithm is used for training the ANN model).

After training the neural network model, we used the model's outputs as predictors of firms' future returns. Neural network model assigned each firm-year into ten portfolios: one (1) being the least favorable and ten (10) being the most favorable. To assess the performance of the artificial neural network, we measured the abnormal return that accrues to the ten portfolios and a hedge portfolio that goes long on the tenth portfolio and short on the first portfolio.

We use the calendar time regression approach (Jensen's alpha) to measure the abnormal performance of the ten portfolios. In the calendar time regression approach we first calculated mean monthly returns for all ten portfolios starting four months after fiscal year end for the duration of one year, we then rebalanced according to the new rankings artificial neural network assigned. This yielded a series of monthly raw returns for each portfolio from April, 1990 to March, 2004. We then computed excess monthly portfolio returns by subtracting the risk free rate. Finally, we regressed (using OLS) the mean excess monthly portfolios returns on the excess CRSP value weighted index, size (SMB), book to market (HML) and momentum factors (UMD), as follows:

$$R_{pt} - R_{ft} = \alpha + \beta_i \left(R_{mt} - R_{ft}\right) + s_i SMB_t + h_i HML_t + u_i UMD_t + \xi_{it}, \tag{4}$$

where R_{pt} is the monthly return for the constructed portfolio on month t, R_{ft} is the Ibbotson One Month Treasury Bill Rate. R_{mt} is the CRSP value weighted index return for month t. SMB_t is the average return on three small market capitalization portfolios minus the average return on three large market capitalization portfolios on month t. HML_t is the average return on two high book-to-market equity portfolios minus the average return on two low book-to-market equity portfolios for month t. UMD_t is the average of the returns on two (big sized and small sized) high prior return portfolios minus the average of the returns on two low prior return portfolios. In this regression, the intercept proxies for the average monthly abnormal return accumulated by holding the portfolio for the estimation period. In the tables we report annualized intercepts.

EMPRICAL RESULTS

We measure the annual abnormal returns that accrue to the ten portfolios formed according to rankings the artificial neural network model generates. Furthermore a hedge portfolio consisting of a long position in the most favorable stocks and a short position in the least favorable stocks is created and its performance is assessed.

In Table 2 we report the abnormal returns that accrue for each portfolio within the validation sample that spans between the years 1990 and 2005. The results suggest that the least favorable portfolio underperforms expected returns generated by the Carhart four-factor model by an average of 28.9% which is statistically significant at the one-percent level. This suggests that shorting least favorable stocks generates a positive abnormal return which is statistically and economically significant. On the other hand the portfolio consisting of the most favorable stocks fails to generate returns that are above the expected returns.

We also test whether a hedge portfolio that goes long on most favorable securities and short on the least favorable ones generates significant abnormal returns. Our results suggest that such a trading strategy generates annual abnormal returns in excess of 22 percent. The hedge portfolio result is consistent with artificial neural network having significant predictive power of future abnormal returns.

Our results, combined, suggest that the artificial neural network model applied in this study has predictive power of future returns. Moreover our results indicate that the predictive power of artificial neural network is not due to previously documented size, book-to-market and momentum anomalies. In other words abnormal returns generated by the neural network trading strategy are in excess of an investor that would follow a combined strategy of previously documented anomalies.

Table : 2 Long-Term Performance of an Analyst Consensus Based Trading Strategy

Portfolio	Annualized Abnormal Return	Intercept	Beta	SMB	HML	UMD	Obs.	R-square
Strong Sell	-28.9%	-0.024*** (-3.66)	1.203*** (6.72)	1.428*** (7.61)	-0.538** (-2.31)	-0.408*** (-3.06)	199	55.1%
2	-13.3%	-0.011 (-0.9)	1.372 (4.12) ***	0.124 (0.35)	-0.953*** (-2.19)	0.032 (0.13)	200	19.9%
3	7.7%	0.006* (1.76)	1.006*** (10.2)	0.555*** (5.38)	0.051 (0.39)	0.077 (1.04)	200	51.6%
4	-3.0%	-0.002 (-0.84)	1.279*** (15.95)	0.063 (0.75)	-0.516*** (-4.93)	-0.103 (-1.71)	200	74.1%
5	3.8%	0.003** (2.63)	0.979*** (29.62)	-0.048 (-1.38)	-0.386*** (-8.94)	-0.055** (-2.23)	200	90.3%
6	-0.3%	0.000 (-0.25)	0.955*** (34.12)	0.006 (0.19)	0.192*** (5.25)	-0.038* (-1.82)	200	88.3%
7	-2.0%	-0.002 (-1.54)	1.013*** (34.28)	0.055* (1.77)	0.459*** (11.89)	-0.032 (-1.44)	200	87.1%
8	-4.6%	-0.004 (-1.11)	1.082*** (11.56)	0.795*** (8.11)	-0.105 (-0.86)	-0.001 (-0.01)	200	64.7%
9	1.3%	0.001 (0.18)	1.109*** (6.65)	0.893*** (5.11)	-0.093 (-0.43)	0.003 (0.02)	200	38.9%
Strong Buy	-3.5%	-0.003 (-0.54)	1.293*** (8.74)	0.430 (2.78) **	0.214 (1.11)	-0.320** (-2.89)	200	41.0%
Hedge Portfolio	22.2%	0.019** (2.26)	-0.068 (-0.31)	-0.833 (-3.56) *	0.857 (2.96) **	0.017 (0.1)	199	19.4%

*This table reports the abnormal performance of portfolios constructed using ratings assigned by the neural network model. Each firm in the database is distributed to one of the ten portfolios (strong sell – strong buy) with respect to its rating. And the eleventh portfolio titled "Hedge" represents the returns to a trading strategy that goes long on the portfolio with the strong buy stocks and short on the strong sell stocks. Using the calendar time portfolio regression approach the Jensen's alpha of each portfolio is computed and reported, with its t-statistics, on the first row of both panels under the label "Intercept (Abnormal Return)". The following row titled beta shows the beta (with t-statistics), SMB, HML and UMD factor sensitivities of all eleven portfolios. Finally the last two rows in each panel report the r-square of the asset pricing model. The first figure in each cell is the regression coefficient. The second figure in each cell is the t-statistic. *, **, and *** indicate at the 10, 5, and 1 percent levels respectively.*

However, as in market efficiency studies, we would like to note that our results are subject to a joint hypothesis problem which restrains us from attributing these to market efficiency or mispricing as these results may be an artifact of an incomplete asset pricing model. Nevertheless, assuming the Carhart four-factor model to be a complete model we find evidence supporting the use of artificial neural network model as a successful tool for predicting future stock returns.

CONCLUSION

The purpose of this paper was to estimate the abnormal returns that can be earned through an investment strategy based on ratings assigned by an artificial neural network model. We find that over the 1990 to 2005 period, a portfolio of the stocks with the most (least) favorable neural network rating provides an average annual abnormal gross return of -3.5 (-28.9) percent, after controlling for market risk, size, book-to-market, and momentum effects. A hedge strategy of purchasing stocks that have the highest neural network ratings and selling short stocks with the lowest ratings generates an abnormal gross return of 1.9 percent a month. The fundamental analysis trading strategy generated by artificial neural networks yields an average annual abnormal return of 22.32%. Our results show neural network's ability to predict future returns in NYSE/AMEX/Nasdaq securities for the period 1990-2005.

REFERENCES

Abarbanell, J.S. and Bushee, B.J. (1997), Fundamental analysis, Future earnings, ans stock prices. *Journal of Accounting Research,* 35, 1-24.

Abarbanell, J.S. and Bushee, B.J. (1998), Abnormal returns to fundamental analysis strategy. *The Accunting Review,* 73, 19-45.

Abdel-khalik, A.R. (1990), Specification problems with information content of earnings: revisions and rationality of expectations and self-selection bias. *Contemprary Accounting Research,* 7, 142-172.

Albanis, G. and Batchelor, R. (2007), Combining heterogeneous classifiers for stock selection. *Intelligent Systems in Accounting, Finance and Management*, 15, 1-21.

Beaver, W., Clarke, R. and Wright, F. (1979), The association between unsystematic security returns and the magnitude of earnings forecast errors. *Journal of Accounting Research,* 17, 316-340.

Bishop, C.M. (1997), *Neural Networks for Pattern Recognition.* Oxford, Clerendon Press.

Carter, R.B. and Van Auken, H.E. (1990), Security analysis and portfolio management: a survey and analysis. *Journal of Portfolio Management*, 16, 81-85

Cheng, B. and Titterington, D.M. (1994), Neural networks:A review from statistical perspective. *Statistical Science ,* 9, 2-54.

Coakley, J.R. and Brown, C.E. (2000), Artificial neural networks in accounting and finance:modeling issues. *International Journal of Intelligent Systems in Accounting, Finance Management ,* 9, 119-144.

Freeman, R. and Tse, S. (1989), The multiperiod information content of accounting earnings:confirmations and contradictions of previous earnings reports. *Journal of Accounting Research*, 27, 49-84.

Goontilake, S. and Treleaven, P. (1995), *Intelligent Systems for Finance and Business.* Chichester, Wiley.

Greig, A.C. (1992), Fundamental analysis and subsequent stock returns. *Journal of Accounting and Economics,* 15, 413-422.

Haykin, S. (1999), *Neural Networks (2nd Edition).* N.J., Prentice Hall.

Hecht-Nilsen, R. (1990), *Neurocomputing.* Massachusetts, Addison-Wesley Pub. Comp.

Hill, T., Marquez, L., O'Connor, M. and Remus, W. (1994), Artificial neural network models for forecasting and decision making. *International Journal of Forecasting,* 10, 177-191.

Holthausen, R.W. and Larcker, D.F. (1992), The prediction of stock returns using financial statement information. *Journal of Accounting and Economics*, 15, 373-411.

Krishnaswamy, C., Gilbert, E.W. and Pashley, M.M. (2000), Neural network applicartios in finance: a practical introduction. *Financial Pracice and Education*, Spring/Summer 75-84.

Kryzanowski, L., Galler, M. and Wright, D.W. (1993), Using artificial neural networks to pick stocks. *Financial Analysts Journal,* 49, 21-27.

Lev, B. and Thiagarajan, R,S. (1993), Fundamental information analysis. *Journal of Accounting Research,* 31, 190-215.

Omran, M. and Ragab, A. (2004), Linear versus non-linear relationships between financial ratios and stock returns:emprial evidence from Egyptian firms. *Review of Accounting and Finance ,* 3: 84-102.

Ou, A. and Penman, S.H. (1989), Financial statement analysis and the prediction of stock returns. *Journal of Accounting and Economics ,* 11, 295-329.

Pahor, M. and Mramor, D. (2001), Testing nonlinear relationships between rate of return on equiy and financial ratios. *Working Paper* .

Quah, T.S. and Srinivasan B. (1999). Improving returns on stock investment through network selection. *Expert Systems with Applications*, 17, 295-301.

Schalkof, R.J. (1997), *Artificial Neural Networks.* New York, The McGraw-Hill Comp.,Inc.

Trippi, R.R. and Efraim, T. (1996), *Neural Network in Finance and Investing.* Chicago, Irwin Professional Pub.

Vellido, A., Lisboa, P. and Vaughan, J. (1999), Neural networks in business:a survey of applications (1992-1998). *Expert Systems and Applications ,* 17, 51-70.

Wong, B.K. and Selvi, Y. (1998), Neural network applications in finance:a review and analysis of literature (1990-1996). *Information Management ,* 34, 129-139.

Wong, B.K., Bodnovich, T.A. and Selvi, Y. (1997), Neural network applications in business:a review and analysis of the literature 1988-95. *Decision Support Systems,* 19, 301-320.

Wong, B.K., Lai, V.S. and Lam, J. (2000), A bibliography of neural network business applications research:1994-1998. *Computers Operatins Research ,* 27, 1045-1076.

Wong, F.S., Wang, P.Z., Goh, T.H. and Quek, B.K. (1992) Fuzzy Neural Systems for Stock Selection. *Financial Analysts Journal*, 48, 47-52.

Zhang, G., Patuwo, B.E. and Hu, Y.M. (1998), Forecasting with artificial neural networks:state of the art. *International Journal of Forecasting ,* 14, 35-62.

ACKKNOWLEDGE

We are indebted to Miklos Vasarhelyi and workshop participants at Rutgers University for their valuable comments.

BIOGRAPHY

Birol Yildiz is an Associate Professor of Accounting and Finance at Eskisehir Osmangazi University. His research focuses on artificial intelligence applications on accounting and financial analysis. He can be contacted at: IIBF, Eskisehir Osmangazi University, 26480, Eskisehir, Turkey. Email: byildiz@ogu.edu.tr

Ari Yezegel is an assistant professor in accounting at Bentley University. His research focuses on accounting regulations, cross-listings, disclosure and financial analysts. Some of Ari's research has appeared in *Journal of Banking and Finance, Financial Analysts Journal* and *Quantitative Finance*. Ari holds a Ph.D. degree in Accounting from Rutgers University. He can be contacted at: 175 Forest Street Waltham, Massachusetts 02452 USA. Email: ayezegel@bentley.edu