

# Predicting Audit Reports Using Meta-Heuristic Algorithms

Hashem Valipour\*, Fatemeh Salehi\*\*, Mostafa Bahrami\*\*\*

## Abstract

**Purpose** - This study aims to predict the audit reports of listed companies on the Tehran Stock Exchange by using meta-heuristic algorithms.

**Research design, data, methodology** - This applied research aims to predict auditors reports' using meta-heuristic methods (i.e., neural networks, the ANFIS, and a genetic algorithm). The sample includes all firms listed on the Tehran Stock Exchange. The research covers the seven years between 2005 and 2011.

**Results** - The results show that the ANFIS model using fuzzy clustering and a least-squares back propagation algorithm has the best performance among the tested models, with an error rate of 4% for incorrect predictions and 96% for correct predictions.

**Conclusion** - A decision tree was used with ten independent variables and one dependent variable the less important variables were removed, leaving only those variables with the greatest effect on auditor opinion (i.e., net-profit-to-sales ratio, current ratio, quick ratio, inventory turnover, collection period, and debt coverage ratio).

**Keywords** : Audit report, ANFIS, Tehran Stock Exchange.

**JEL Classifications** : C61, C63, C67.

## 1. Introduction

Human being has always been dealing with the concept of "decision-making", from their early days to now that they have entered today's complex, turbulent business environment. Meanwhile, experts in management like Herbert Simon go so far as to claim that management is all about rational decision-making. Decision-making is a selection problem. In such problems, decision parameters are very complex and in some cases infeasible. On the other hand, access to relevant and reliable information can reduce biases in decision-making and help the users of financial information in making optimal

decisions. Auditing has made such information available, but improving the efficiency of auditing has always been an issue for auditors and users.

The rationale behind auditing is to accredit the information reflected in financial statements, for this information forms the basis for decision-making by different groups such as shareholders, potential investors, agents, managers, financial consultants, analysts, creditors, and the government. From the view of these users, auditing is effective when auditors verify that lack of any errors, faults, fraud, and manipulation in financial statements.

The development of new technologies and their application in different sciences have also affected accounting and auditing. One way of increasing audit efficiency is to use auditor's opinion forecast techniques in order to develop a suitable model for auditing conditions and environment.

When the number of input variables to a system and the relationships between them increase exponentially, the forecast results will show more biases than real outputs. There are various ways for analyzing these relationships, the most robust of which is artificial neural networks (ANN).

Artificial intelligence methods have recently been used in the prediction and optimization of decision-making in financial market. These methods can be used to maximize profit and minimize investment risk in unpredictable environments. Artificial neural networks, neuro-fuzzy networks, and meta-heuristic algorithms are among these dynamic systems. The main advantage of these intelligent systems is their application in complex non-linear functions.

The present research examines the factors that affect audit reports and the possibility of predicting audit reports using meta-heuristic methods. Therefore, the accounting data of the firms listed on Tehran Stock Exchange (TSE) is examined using adaptive neuro-fuzzy inferences system (ANFIS) to see whether this method can be used in predicting auditor's opinion.

## 2. Review of the Literature

Sajjadi (2008) applied logistic regression and chi-squared test to examine the factors that affect qualified audit report. The results suggested that there are significant relationships between audit report and firm size, current ratio, debt-to-asset ratio, accounts receivable to total assets ratio, and audit firm type.

Using artificial neural networks, Ahmadpour et al. (2010) studied the effect of financial and non-financial variables on whether a quali-

\* First Author, Assistant Professor of Accounting, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran

\*\* Corresponding Author, Kish International Branch, Islamic Azad University, Kish Island, Iran, E-mail: [salehi19862@yahoo.com](mailto:salehi19862@yahoo.com).

\*\*\* Hidaj Branch, Islamic Azad University, Hidaj, Iran

fied opinion report is issued. The results showed that variables such as quick ratio, inventory turnover, total debt to total assets ratio, and accounts receivable to total assets ratio had significant effects on issuing a qualified opinion report.

Saif et al. (2012) investigated the possibility of predicting the type of audit report using a combination of a support vector machine and a decision tree. Studying a sample of TSE-listed firms (708 firms with qualified and 310 firms with unqualified opinion), the results showed that the algorithm is a powerful tool for predicting qualified and unqualified opinion reports.

Kirkos et al. (2008) studied the variables that affect auditor appointment in a sample of 338 UK and Irish firms using three data mining classification techniques (decision trees, neural networks, and support vector machines). The results showed that the total accuracy of the models were 83.73% for the decision tree, 79.29% for SVM, and 75.44% for the neural network model. These models suggested that the level of debt is a factor that influences the auditor choice decision.

Doumpos et al. (2005) examine whether SVM models can predict the type of audit report. The results showed that all SVM models were capable of distinguishing between qualified and unqualified financial statements with satisfactory accuracy. However, nonlinear SVM models were less accurate than linear SVM models in predicting paragraphs in qualified opinion reports.

Lenard et al. (1995) used a neural network and a log it model to predict which firms would receive audit reports reflecting a going concern uncertainty modification. The results showed that the neural network model is more accurate than the log it model with an accuracy of 95%.

Spa this (2003) used logistic regression to detect falsified financial statements. The reported accuracy rate exceeded 84%. The results suggested that there is potential in detecting FFS through the analysis of published financial statements.

Gaganis et al. (2007) applied probabilistic neural networks to the identification of qualified audit opinions. The results suggested the high explanatory power of neural network model, which outperformed the multi layer perceptron model and the logistic model.

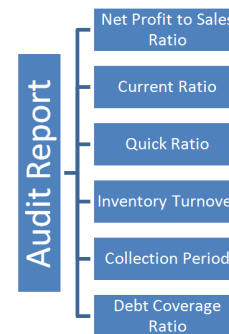
Kirkos et al. (2007) employed three data mining classification techniques (i.e. decision tree, multi layer perceptron neural network, and Bayesian belief network) to identify qualified auditor's opinion and its determinants. The results suggested the greater accuracy of the decision tree model against the training set. However, based on ten-fold cross validation results, the Bayesian belief network achieved the highest accuracy in identifying the cases of qualified audit opinions.

The key to success in today's competitive world is to reduce time and costs and to increase efficiency. Competition has also substantially increased in the area of auditing across the world, especially in advanced countries (Firth, 1993). This competitive environment has forced audit firms to find ways of increasing their efficiency. The purpose of the present research is to examine the ability of meta-heuristic methods such as ANFIS to predict auditor's opinion. Using the outputs of neural networks and meta-heuristic algorithms, we can identify the most probable auditor's opinion before audit takes place. Also the most efficient model can be used to develop an auditing

process for achieving an acceptable level of audit risk. The main question of the research is whether ANFIS is able to predict auditor's opinion.

### 3. Conceptual Model

The factors that affect audit opinions are as follows:



<Figure 1> The conceptual model of the research

### 3.1. Methodology

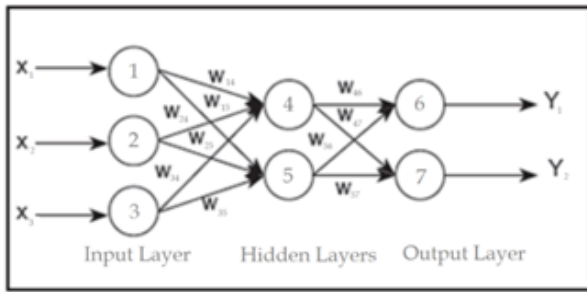
#### 3.1.1. Population and sample

The present research is an applied research which aims to predict auditor's report using meta-heuristic methods (i.e. neural networks, ANFIS, and genetic algorithm). The population consists of all the firms listed in Tehran Stock Exchange (TSE). The research covers a 7-year period from 2005 to 2011. The sample is selected based on the following four criteria:

1. Firms that are not financial intermediaries;
2. Firms whose financial year ends at the end of Iranian calendar;
3. Firms with audit report for at least one year in the research period;
4. Firms with full financial information, including annual financial statements, audit report, and notes for the period in which the firm is studied.

#### 3.1.2. Artificial neural networks

A neural network consists of a number of interconnected processing units called neurons, with each neuron having inputs and outputs and performing a relatively simple local task. Artificial neural network (ANN) is a simplified model of the central nervous system that mimics the structure of the human brain. It involves a network of simple processing elements exhibiting complex global behavior determined by the connections between the processing elements and element parameters. One of the most important features of neural networks is their learning ability. Neural networks generally consist of an input layer, one or several hidden layers, and an output layer (Kartalopoulos, 1996). An ANN is shown in Figure 2.



<Figure 2> An example of an artificial neural network

The general concepts in ANN are: artificial neurons, activation function, network learning, and back propagation learning algorithm.

### 3.1.3. Genetic algorithm

Genetic algorithm is an optimization method that mimics the process of natural evolution. It can be classified as a numerical, direct, and stochastic method. This algorithm is based on iteration where a population of candidate solutions to an optimization problem is evolved toward better solutions. Genetic algorithm can be applied to various problems such as optimization, system identification and control, image processing, topology identification, and training of artificial neural networks and decision-based systems.

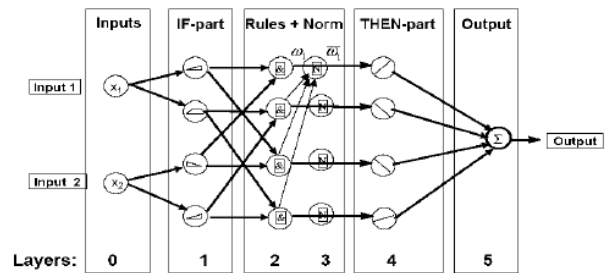
Genetics is the science of genes, heredity, and variation in living organisms. The main factors in inheritance of biological characteristics are chromosomes and genes. Genetic algorithm (GA), which is based on these ideas, is used for optimization, search, and machine learning. It is based on Charles Darwin's principle of "survival of the fittest", where the strongest genes and chromosomes dominate the weaker ones. In fact, evolution occurs on the encoded strings rather than the living organisms representing the strings. Based on the idea of natural selection, the fittest the organism, the more likely it is to survive and reproduce. In GA, each candidate solution is represented by an array of bits. The implementation of GA involves four basic steps: (1) initialization, where an initial population of solutions is randomly generated, (2) selection, where individual solutions are selected through a fitness-based process, (3) crossover, which represents mating between parent solutions to produce a child solution, and (4) mutation, where random modifications are introduced.

Through crossover and mutation, GA ensures that the search space is adequately searched, thus reducing the chance of convergence towards local optima. Selection and crossover distinguish genetic algorithms from a purely random search and direct the algorithm toward finding an optimum. Mutation is designed to ensure diversity in the search to prevent premature convergence to a local optimum.

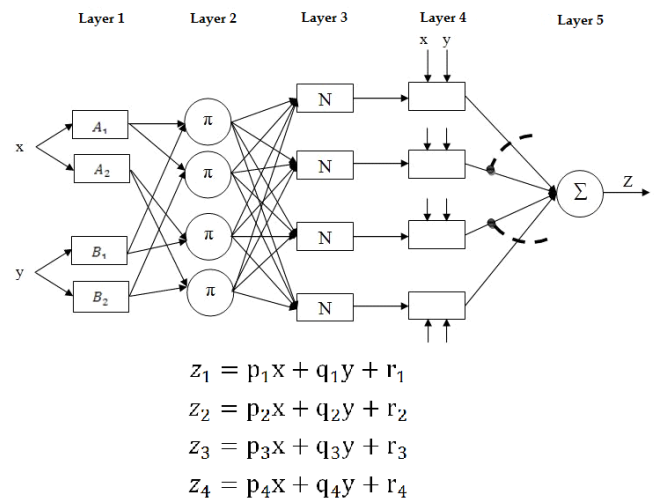
### 3.1.4. Adaptive neuro-fuzzy inference system (ANFIS)

ANFIS integrates the principles of neural networks and fuzzy logic to map an input to an output space. Its basic mechanism is a set of if-then rules. ANFIS utilizes linguistic information from the fuzzy logic as well learning capability of an ANN for automatic fuzzy if-then rule generation and parameter optimization. It was first in-

troduced by Jang (1993). ANFIS is a multi layer network consisting of nodes and directional links. The structure of ANFIS is schematically illustrated in Figure 3.



<Figure 3> The structure of ANFIS



<Figure 4> An ANFIS model with two input variables

As can be seen in the above figures, the structure of ANFIS consists of five layers:

- Layer 1: Each node output in this layer is fuzzified by membership grade of a fuzzy set corresponding to each input.
- Layer 2: Each node output in this layer represents the firing strength of a rule, which performs fuzzy, AND operation.
- Layer 3: Each node output in this layer is the normalized value of layer 2, i.e., the normalized firing strengths.
- Layer 4: Each node output in this layer is the normalized value of each fuzzy rule. The nodes in this layer are adaptive.
- Layer 5: The node output in this layer is the overall output of the system, which is the summation of all coming signals.

An important issue in any ANFIS model is the type of membership function that suits the model's inputs. There are different types of membership functions, such as bell, Gaussian, trapezoidal, triangular, sigmoidal, and Pi.

### 3.2. Variable selection

The variables used in the present research are listed in Table 1.

<Table 1> The variables

Variable	Abbreviation	Measurement
Inventory Turnover	X1	
Current Ratio	X2	Current assets divided by current liabilities
Quick Ratio	X3	Current assets less inventories, divided by current liabilities
Gross Profit to Sales	X4	
Collection Period	X5	
Net Profit to Sales	X6	
Working Capital Turnover	X7	
Total Asset Turnover	X8	
Debt Coverage Ratio	X9	
Return on Assets	X10	Net income divided by average total assets

In this research, we created a decision tree with 10 independent variables and 1 dependent variable. Based on relative weight, the most important independent variables were selected as factors that have the greatest effect on auditor’s opinion (Table 2).

<Table 2>The most important independent variables

Independent Variable	Abbreviation
Net Profit to Sales	X6
Current Ratio	X2
Quick Ratio	X3
Inventory Turnover	X1
Collection Period	X5
Debt Coverage Ratio	X9

### 3.2.1. The ANFIS model

Fuzzy systems have a special place in modern modeling methods. ANFIS has evolved from the combination of fuzzy systems which are based on logical rules and neural networks which are able to extract knowledge from numerical data (Nayak et al., 2003). ANFIS maps the relationships between inputs and outputs using the learning capacity of neural networks and yields the results using fuzzy logic.

Since multi layer perceptron artificial neural networks randomly consider their weight values, they may get stuck in local minima; thus, the network will not be trained well and will not be efficient. ANFIS, however, has fewer constraints than ANNs and has various applications in modeling, decision-making, and signal processing and control.

## 4. Results

### 4.1. Subtractive clustering and back propagation (BP) algorithm

<Table 3> The results of ANFIS modeling with subtractive clustering and BP algorithm

Data Sets	Percentage	N	Error
Training	80%	130	0.07665439
Checking	20%	33	0.0476869

<Table 4> The results of ANFIS prediction with subtractive clustering and BP algorithm

Data Sets	Percentage	N	Error
Training	80%	130	0.0531935
Checking	20%	33	0.0844276
Total	100%	163	0.06881055

The table below shows the results of predicting auditor’s opinion with 33 checking samples.

<Table 5> Prediction of auditor’s opinion using ANFIS with subtractive clustering and BP algorithm

Prediction of Auditor’s Opinion	Total	Unqualified	Qualified
Correct prediction	0.915572	0.9236137	0.9075311
Incorrect prediction	0.084428	0.0763863	0.0924689
Number of correct predictions	30	11	19
Number of incorrect predictions	3	1	2

The results suggest the satisfactory performance of the ANFIS model in identifying types of auditor’s opinion with respect to the testing data set. The error rate is 8% for incorrect predictions and 92% for correct predictions. Based on the checking data set for the year 2011 which included 33 opinions, 21 opinions were qualified and 12 opinions were unqualified. This indicates that 2 cases of qualified opinion and 1 case of unqualified opinion were incorrectly predicted, which is a desirable result.

### 4.2. Subtractive clustering and Hybrid LS-BP algorithm

The ANFIS model with subtractive clustering and hybrid least-squares back propagation (LS-BP) algorithm has 10 inputs and 1 output. This model has 163 data for the variables of one year and an output for the auditor’s opinion of the next year. We found that the ANFIS model had a high accuracy.

<Table 6> The results of ANFIS modeling with subtractive clustering and hybrid LS-BP algorithm

Data Sets	Percentage	N	Error
Training	80%	130	0.067809
Checking	20%	33	0.043456

<Table 7> The results of ANFIS prediction with subtractive clustering and LS-BP algorithm

Data Sets	Percentage	N	Error
Training	80%	130	0.030941837
Checking	20%	33	0.044142378
Total	100%	163	0.037542107

<Table 8> Prediction of auditor’s opinion using ANFIS with subtractive clustering and LS-BP algorithm (33 checking samples)

Prediction of Auditor’s Opinion	Total	Unqualified	Qualified
Correct prediction	0.95585762	0.93158064	0.9801346
Incorrect prediction	0.04414238	0.06841936	0.0198654
Number of correct predictions	32	11	21
Number of incorrect predictions	1	1	0

The results of prediction using ANFIS with subtractive clustering and hybrid LS-BP algorithm suggests the satisfactory performance of the model in identifying types of auditor’s opinion with respect to the testing data set. The error rate is 4% for incorrect predictions and 96% for correct predictions. Based on the checking data set for the year 2011 which included 33 opinions, 21 opinions were qualified and 12 opinions were unqualified. This indicates that all the cases of qualified opinion were predicted correctly, while only 1 case of unqualified opinion was incorrectly predicted.

#### 4.3. FCM and BP algorithm

The ANFIS model with fuzzy clustering (FCM) and BP algorithm has 10 inputs and 1 output. This model has 163 data for the variables of one year and an output for the auditor’s opinion of the next year. The results suggested that the ANFIS model had a high accuracy.

#### 4.4. FCM and LS-BP algorithm

The ANFIS model with FCM and hybrid LS-BP algorithm has 10 inputs and 1 output. This model has 163 data for the variables of one year and an output for the auditor’s opinion of the next year. The results suggested that this model predicted the majority of firms correctly.

<Table 9> The results of ANFIS modeling with FCM and hybrid LS-BP algorithm

Data Sets	Percentage	N	Error
Training	80%	130	0.03837851
Checking	20%	33	0.04577278

<Table 10> The results of ANFIS prediction with FCM and LS-BP algorithm

Data Sets	Percentage	N	Error
Training	80%	130	3.11311E-05
Checking	20%	33	0.044873767
Total	100%	163	0.022452449

<Table 11> Prediction of auditor’s opinion using ANFIS with FCM and LS-BP algorithm (33 checking samples)

Prediction of Auditor’s Opinion	Total	Unqualified	Qualified
Correct prediction	0.955126	0.95789968	0.95235279
Incorrect prediction	0.044874	0.04210032	0.04764721
Number of correct predictions	31	11	20
Number of incorrect predictions	2	1	1

The ANFIS model with FCM and hybrid LS-BP algorithm has a satisfactory performance in identifying types of auditor’s opinion with respect to the testing data set. The error rate is 4% for incorrect prediction and 96% for correct prediction. Based on the checking data set for the year 2011 which included 33 opinions, 21 opinions were qualified and 12 opinions were unqualified. This indicates that one case of qualified opinion and one case of unqualified opinion were incorrectly predicted.

#### 4.5. Model comparisons

There sults in the table below show that the ANFIS model with subtractive clustering and hybrid LS-BP algorithm has the highest accuracy.

<Table 12> Model comparisons

Model	Prediction of Auditor’s Opinion	Total	Unqualified	Qualified
Subtractive clustering and BP algorithm	Correct prediction	0.915572	0.9236137	0.9075311
	Incorrect prediction	0.084428	0.0763863	0.0924689
Subtractive clustering and LS-BP algorithm	Correct prediction	0.95585762	0.93158064	0.9801346
	Incorrect prediction	0.04414238	0.06841936	0.0198654
FCM and BP algorithm	Correct prediction	0.933636	0.92601153	0.9412606
	Incorrect prediction	0.066364	0.07398847	0.0587394
FCM and LS-BP algorithm	Correct prediction	0.955126	0.95789968	0.95235279
	Incorrect prediction	0.044874	0.04210032	0.04764721

## 5. Conclusion

This research examined whether adaptive neuro-fuzzy inference systems (ANFIS) are capable of predicting auditor's opinion. It also compared different ANFIS models. The results supported the main hypothesis.

The 2011 data of the firms were extracted for measuring the independent variables. Using a decision tree with 10 independent variables and 1 dependent variable, the less important variables were removed, leaving only those variables that have the greatest effect on auditor's opinion—i.e. net profit to sales ratio, current ratio, quick ratio, inventory turnover, collection period, and debt coverage ratio.

The results showed that the ANFIS model with fuzzy clustering and least-squares back propagation algorithm has the best performance among the tested models. Its error rate was 4% for incorrect predictions and 96% for correct predictions.

### 5.1. Limitations

1. Since auditor's opinion is one of the most private information of firms, the researcher was forced to study TSE-listed firms (disclosure of financial information is one of the requirements of listing firms on TSE).

2. Due to unavailability of non-financial information such as firm size and previous audit results, these variables could not be accounted for in prediction of auditor's opinion.

Received: May 02, 2013.

Revised: May 20, 2013.

Accepted: June 17, 2013.

## References

- Ahmadpour, A., Taherabadi, A. A. & Abbasi S. (2010), "The Effect of Financial and Non-Financial Variables on the Issuance of Qualified Auditor's Opinion", *Journal of Tehran Stock Exchange*, 3, 97-114.
- Bäck, T. & Hoffmeister, F. (1991), "Extended Selection Mechanisms in Genetic Algorithms", *Proceedings of the 4th IEEE International Conference on Genetic Algorithms*, 89-99.
- Bingul, Z., Sekmen, A. & Zein-Sabatto, S. (2000), "Evolutionary Approach to Multi-Objective Problems Using Adaptive Genetic Algorithms", *2000 IEEE International Conference on Systems, Man, and Cybernetics*, 3, 1923-1927.
- Chakraborty, M. & Chakraborty, U. K. (1997), "An Analysis of Linear Ranking and Binary Tournament Selection in Genetic Algorithms", *Proceedings of International Conference on Information, Communications and Signal Processing*, 1, 407-411.
- Charbonneau, P. (1997), *An Introduction to Genetic Algorithms for Numerical Optimization*, Boulder, Colorado: National Center for Atmospheric Research.
- Coats, P. K. & Fant, L. F. (1991), "A Neural Network Approach to Forecasting Financial Distress", *Journal of Business Forecasting*, 10, 9-12.
- Coit, D. W. & Smith, A. E. (1999), "Using a Neural Network as a Function Evaluator during GA Search for Reliability Optimization", *Department of Industrial Engineering, University of Pittsburgh*.
- Doumpos, M., Gaganis, C. & Pasiouras, F. (2005), "Explain Qualifications in Audit Reports Using a Support Vector Machine Methodology", *Intelligent Systems in Accounting, Finance and Management*, 13, 197-215.
- Edelstein, H. A. (1999), *Introduction to Data Mining and Knowledge Discovery*, Third Edition, Potomac, MD: Two Crows Corporation.
- Firth, M. (1993), "Price Setting and the Value of a Strong Brand Name", *International Journal of Research in Marketing*, 10, 381-386.
- Gaganis, C., Pasiouras, F. & Doumpos, M. (2007) "Probabilistic Neural Networks for the Identification of Qualified Audit Opinions", *Expert Systems with Applications*, 32, 114-124.
- Haldenbilen, S. & Ceylan, H. (2005), "Genetic Algorithm Approach to Estimate Transport Energy Demand in Turkey", *Energy Policy*, 33, 89-98.
- Jang, J. S. R. (1993), "ANFIS: Adaptive Network-Based Fuzzy Inference Systems", *IEEE Transactions on Systems, Man and Cybernetics*, 23, 665-685.
- Kartalopoulos, S.V. (1996), *Understanding Neural Networks and Fuzzy Logic: Basic Concepts and Applications*, New York, NY: the Institute of Electrical and Electronics Engineers, Inc.
- Kato, K., Skinner, D. J. & Kunimura Sr., M. (2009), "Management Forecasts in Japan: An Empirical Study of Forecasts that are Effectively Mandated", *Accounting Review*, 84, 26-38.
- Kennedy J. & Eberhart, R. C. (1995), "Particle Swarm Optimization", *Proceedings of IEEE International Conference on Neural Networks*, 4, 1942-1948.
- Kirkos, E., Spathis, C., Nanopoulos, A. & Manolopoulos, Y. (2007) "Identifying Qualified Auditors' Opinions: A Data Mining Approach", *Journal of Emerging Technologies in Accounting*, 4, 183-197.
- Kirkos, E., Spathis, C. & Manolopoulos, Y. (2008), "Support Vector Machines, Decision Trees and Neural Networks for Auditor Selection", *Journal of Computational Methods in Sciences and Engineering*, 8, 213-224.
- Kohzadi, N., Boyd, M. S., Kermanshahi, B. & Kaastra, I. (1975), "A Comparison of Artificial Neural Network and Time Series Models for Forecasting Commodity Prices", *Neurocomputing*, 10, 169-181.
- Lenard, M., Alam, P. & Madey G. (1995), "The Application of Neural Networks and a Qualitative Response Model to the Auditor's Going Concern Uncertainty Decision", *Decision Sciences*, 26, 209-227.
- Lintner, J. (1956), "Distribution of Income of Corporations among Dividends, Retained Earnings and Taxes", *American Economic Review*, 46, 97-113.

- Medeiros, M. C. (2000), "A Hybrid Linear-Neural Model for Time Series Forecasting", *IEEE Transactions on Neural Networks*, 11, 1402-1412.
- Nayak, P. C., Sudheer, K. P., Rangan, D. M. & Ramasastri, K. S. (2004), "A Neuro-Fuzzy Computing Technique for Modeling Hydrological Time Series", *Journal of Hydrology*, 291, 52-66.
- Saif, M., Sarikhani, M. & Ebrahimi, F. (2012), "Finding Rules for Audit Opinions Prediction through Data Mining Methods", Working Paper.
- Sajjadi, H., Farazmand, H., Dastgir, M. & Dehghanfar, D. (2007), "The Driving Factors of Qualified Opinion Report", *Iranian Journal of Accounting Studies*, 18, 123-146.
- Shi, Y. (1998), "A Modified Particle Swarm Optimizer", Proceedings of the 1998 IEEE International Conference on Evolutionary Computation, 69-73.
- Spathis, C. (2002), "Detecting False Financial Statements Using Published Data: Some Evidence from Greece", *Managerial Auditing Journal*, 17, 179-191.