International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:9, No:7, 2015

# Detecting Earnings Management via Statistical and Neural Network Techniques

Mohammad Namazi, Mohammad Sadeghzadeh Maharluie

***Abstract*—**Predicting earnings management is vital for the capital market participants, financial analysts and managers. The aim of this research is attempting to respond to this query: Is there a significant difference between the regression model and neural networks' models in predicting earnings management, and which one leads to a superior prediction of it? In approaching this question, a Linear Regression (LR) model was compared with two neural networks including Multi-Layer Perceptron (MLP), and Generalized Regression Neural Network (GRNN). The population of this study includes 94 listed companies in Tehran Stock Exchange (TSE) market from 2003 to 2011. After the results of all models were acquired, ANOVA was exerted to test the hypotheses. In general, the summary of statistical results showed that the precision of GRNN did not exhibit a significant difference in comparison with MLP. In addition, the mean square error of the MLP and GRNN showed a significant difference with the multi variable LR model. These findings support the notion of nonlinear behavior of the earnings management. Therefore, it is more appropriate for capital market participants to analyze earnings management based upon neural networks techniques, and not to adopt linear regression models.

***Keywords*—**Earnings management, generalized regression neural networks, linear regression, multi-layer perceptron, Tehran stock exchange.

## I. Introduction

EARNINGS management is a popular and appealing concept in contemporary management accounting research, and is considered as one of the most significant current topic in accounting [1]-[3]. It encompasses various concepts such as: signaling private information by management, opportunistic choices within the GAAP restriction, and committing mistakes and frauds. Understanding different methods of earnings management and the reasons behind its exertion can help standard setting boards in promulgating or changing related standards. In addition, earnings management has recently been investigated as one of the key subjects in investigating the success of deploying international accounting standards in the world [4]. Also, behind the scene of recent scandals such as Xerox and WorldCom, the role of earnings management cannot be ignored. Given these reasons, accounting researchers have widely started to investigate this phenomenon from the middle of 1980s [4], [5]. After changing accounting standards in the 1990s and financial and accounting scandals, researchers' interest in earnings management's topic has also increased [4].

A major query which arises here is: How earnings management can be detected? Traditional approaches for investigating earnings management, have mainly applied the Linear Regression (LR) model [6], [7]. However, this technique posits several limitations, including linearity and other predetermined model and regression assumptions such as nonexistence of correlation, homoscedasticity and non-flexibility of the regression models [8]. In addition, the existing evidence [6], [9] unambiguously demonstrates that firms' earnings can be characterized in terms of non –linear relations. Thus, Artificial Neural Network (ANN) techniques can be employed in these situations.

A subtle question that arises here, however, is whether ANN techniques posit a superior performance in detecting earnings management than the LR model? The major aim of this research is responding to the preceding question. In order to provide empirical data, information relating to the firms listed in Tehran Stock Exchange (TSE) market will be provided.

The investigation of this subject is important and relevant at least for four reasons. First, in the light of the previous financial scandals, earnings management has become a pivotal international issue in the discussions surrounding the credibility of the financial markets. Second, as [2] indicates, there is a broad interest and implications in the findings of this literature. Third, the debates concerning earnings management detection was and still is an important topic of day-out impact of the accounting standards. Fourth, prediction of earnings management for the financial statements users in the process of evaluation of recent economic performance, predicting future income and evaluation of the value of companies, provides a high priority and importance [3]. The originality of the paper is related to the fact that, for the first time, it compares the prediction power of the LR model with ANN models in an emerging stock market- TSE.

This paper is organized as follows: In section II literature review and hypotheses would be provided. The research methodology would be explained in Section III. Sections IV and V will describe the research models and the findings respectively. The conclusion and discussion and suggestions would be presented in Sections VI and VII respectively.

## II. Literature Review and Hypotheses

To date, an exact definition, which comprises all aspects of the earning management, does not exist. Professional accounting bodies also have not provided a clear definition of the earnings management in their statements, standards or

M. Namazi is Full Professor of Accounting, Shiraz University, Iran (corresponding author Tel: +989171184375 e-mail: mnamazi@rose.shirazu.ac.ir).

M. Sadeghzadeh Maharluie is a PH.D. Candidate of Accounting, Shiraz University, Iran (e-mail: mohammad_sadeghzadeh2000@yahoo.com).

guidelines. In Iran, lack of a clear definition of earnings management and income smoothing is also apparent by the regulatory agencies and authorities of the TSE market [10].

In spite of this situation, researchers have attempted to deal with some aspects of the earnings management. For example, [11] have classified earnings management into three groups: white, gray, and black. White earnings management increases clearness of the financial statements. Black earnings management includes misstatement of the financial statements and fraud. Gray earnings management adjusts the financial statements in the domain of the accounting standards which can be opportunistic or it can increase the efficiency. Scott and Tsai and Chiou have identified the following techniques for earnings management [9], [12]:

1) Revenue recognition,
2) Change in accounting procedures,
3) Timing of new accounting standards implementation,
4) Buying or selling the asset, and
5) Discretionary accruals.

Tsai and Chiou [9] maintained that among the preceding earnings management tools, accounting accruals, especially discretionary accruals, is easier and more susceptible to be expended for the model. Therefore, most of the earnings management researches (for example, [6], [13]) have focused on the discretionary accruals. Hence this study will also utilize it in constructing earnings management model.

There are two ways for calculating total accruals: 1) the balance sheet approach, and 2) the cash flow approach. In the balance sheet approach, total accruals are determined by subtracting depreciation from changes of the working capital, as [6]:

$$TACC_t = \Delta CA - \Delta CL - \Delta Cash - \Delta STDEBT - DEP$$

where: ΔCA = change in the current asset; ΔCL = change in the current debt; Δ Cash = change in the cash and equivalents; ΔSTDEBT = change in the short term debt; DEP= depreciation and depletion expenses.

In the second method (cash flow approach), the cash flow from operation is deducted from the income before extra-ordinary items, as [6]

$$TACC_t = EXBI - CFO$$

From these two approaches, the cash flow approach is more suitable because accruals are sensitive to merge, consolidation, and exchange transaction [13].

Healy and Wahlen examined the literature of earnings management and its application in the standard setting. They specifically investigated empirical evidences in which accruals were expended for earnings management. Their results showed that only specific accruals were exerted for managing earnings [14]. Bergstresser and Philipponb also investigated the motivation for earnings management [15]. Their research results showed that companies whose management compensation contracts were depended on the stock value and stock options, used accruals for changing reported earnings. They also found that in the years with high accruals, management issued more stock options and sold more stock volumes. Namazi and Khansalar investigated the income smoothing in TSE companies with high income and low ratio of market value to book value, versus low income and high ratio of market value to book value. The research period was from 2003 to 2007 and the Jones' model was employed. The research result showed that companies which had a low income and high market value to book value ratio had more motivation to expend discretionary accruals than other companies. For supporting the finding, they also used the Ikel index. The results were the same as Jones' model [16].

With respect to the prediction of the earnings, Tsai and Chiou employed neural networks models to predict the level of earnings management for decreasing the risk of financial distress caused by earnings management, and also for preventing investors from the great loss in the capital market. Their population of the study was Taiwanese Listed Companies. They used eleven factors which affected earnings management and built a model which had a precision prediction in the cases of manipulating earnings upwards with 81% rate. They found that neural network techniques were useful in detecting earnings management [9]. Hoglund also investigated the earnings management in manufacturing firms from 2006 to 2007 in the United States by using regression models and neural network models. The researcher used two statistical models including linear regression and piecewise regression and three neural networks models including MLP[1], SOM[2], and GRNN[3]. The results showed the superior power of the GRNN in detecting earnings management, MLP and piecewise linear regression were in the next level, and linear regression and SOM showed the lowest power in prediction of earnings management among all examined models [6]. Although [6] and [9], among others, have generally shown evidence in favor of the superior performance of the artificial neural networks, but researchers such as [17]-[19] have concluded that neural networks' performance were not superior in comparison with regression models in all instances. These researchers have shown that when models were designed accurately based upon regression models' principles, neural networks' performance were worse or equal with them.

While these studies have been conducted in other countries, and not in Iran, we would like to test the following hypotheses to see whether the preceding findings would hold true in a developing country (TSE market) or not? The premises of these hypotheses are also explained further in Section IV:

H1. The precision of Generalized Regression Neural Networks (GRNN) is greater than Multi-Layer Perceptron Neural Networks (MLP) in predicting earnings management.

H2. The precision of Multi-Layer Perceptron Neural Networks

[1] This neural networks model is in a supervised learning methods classification models.

[2] It is a kind of artificial neural networks which uses an unsupervised learning algorithm.

[3] GRNN like MLP adopt a supervised learning algorithm but the structure is different from MLP.

(MLP) is greater than Linear Regression (LR) in predicting earnings management.

H3. The precision of Generalized Regression Neural Networks (GRNN) is greater than Linear Regression (LR) in predicting earnings management.

## III. Research Methodology

This study examines detecting earnings management via LR and ANN models in the TSE market empirically. Therefore, it is a quantitative research in the domain of the positive studies on the basis of the historical data, and utilized a one-way quasi-experimental research plan. The data is mainly derived from the firms' audited financial statements of the TSE, and Tadbir Pardaz software. Theoretical information was gathered from the Persian and English literatures.

### *A. Population and Sample of the Research*

The population of this study encompassed all related TSE listed companies from 2003 to 2011. TSE is Iran's largest stock exchange and was opened in February 1967. During its first year of activity, only six companies were listed. Then government bonds and certain state-backed certificates were traded in that market. Today TSE has evolved into an exciting and growing marketplace where individuals and institutional investors trade securities of over 342 companies in 39 industries. In fact, TSE, which is a founding member of the Federation of Euro-Asian Stock Exchanges, has been one of the world's best performing stock exchanges in the years from 2002 to 2011.It is also an emerging stock market [20].

A purposive sampling, however, was exerted and companies which had the following criteria were selected:

1. It must have been listed before the end of 2002 in the TSE.
2. Selected companies must have not been part of the investment and financial companies and also in extracting the oil and agriculture activities.
3. The financial year of the companies must have ended to the last day of the Iranian calendar.
4. Selected companies must have not changed the financial year during the designated period.
5. The required information for measuring discretionary accruals for the investigation period must have been available.

After exerting these criteria, the sample study was reduced to 94 companies.

### *B. Research Variables*

Research variables were divided into three groups: dependent, independent and control variables.

### *C. Dependent Variable*

In this study, following many researchers in this area (for example: [6], [13]), the dependent variable, for both regression and neural networks models, was total accruals. Consequently cash flow approach was adopted for calculating total accruals, and then total accruals were calculated by subtracting cash flow from operation from the income before extra-ordinary items as [6]:

$$TACC_t = EXBI - CFO \tag{1}$$

### *D. Independent Variables*

In this research, based upon the studies of [6], [13], and [21]-[23] among others, three independent variables were considered: the reverse of the last year total asset, change in revenue, and total non-current assets. It was expected that a change in revenue would explain current accruals, and non-current assets would control non-current accruals. Consequently calculation of the variables was attempted as follow:

Change in revenue: the difference in revenue of year t and year t-1.

$$\Delta REV_t = REV_t - REV_{t-1} \tag{2}$$

Non-current assets: the total of non-current assets of the company i in year t

### *E. Control Variables*

In order to control the effect of other variables affecting companies' performance, which were not captured by the independent variables, cash flow from operation was used as a control variable.

Also, based upon the studies of [6], [13], [21]-[23], all variables in the equation were divided to total assets of last year (t-1) to control for the firms size.

## IV. Research Model

In this part, regression model was compared with neural networks models in order to examine their prediction effects in earnings managements.

### *A. Regression Model*

The regression model, which was used in this research, was adopted from Jones (1991). This form of Jones' model first suggested by [21] and then assimilated by the researchers such as [6], [13], and [23]-[24], among others. Consequently the following equation was derived:

$$\frac{TACC}{TA_{t-1}} = \beta_0 \frac{1}{TA_{t-1}} + \beta_1 \frac{\Delta REV_t}{TA_{t-1}} + \beta_2 \frac{PPE_t}{TA_{t-1}} + \beta_3 \frac{CFO_t}{TA_{t-1}} \tag{3}$$

DeTienne et al. point out the following limitations of the linear regression models, and defend the application of the neural networks to eliminate these shortcomings [8]:

1) Linear nature of the regression: A major disadvantage of the regression analysis is that it does not posit any index to show whether the linear style of the data are the best instance. Considering the nature of the social science, the linear models of the statistic is unsuitable.
2) Predetermined models: When regression models are used, the basic model must have been predetermined [8]. This leads to solving the problem easily, but in addition there is a need to guess more.

3) Regression assumptions: The performance of the regression models, depend on the various assumptions such as non-existence of the correlation and normal distribution of the residuals.
4) Non flexibility: Multi variable regressions do not maintain flexibility when one cannot guess the elements of the model.

*B. Neural Networks Models*

Neural networks models eliminate the preceding obstacles, and some empirical studies (for example, [6], and [9]) have shown the advantages of neural networks techniques over LR models in examining earnings management issues. Hence, in this study, the following neural networks models were applied:

1. Multi-Layer Perceptron (MLP)

This neural networks model is in a supervised learning methods classification, and therefore is suitable for substitution with regression models [6]. MLP is among the widely used neural networks models. In applying it, choosing the right amount of the layers and inputs can approximate a nonlinear mapping with a suitable precision. The structure of MLP includes an input layer, one or more hidden layer and an output layer. Each layer comprises of one or more neurons. Input layer has the same neuron as the number of the independent variables. Also, the output layer has the same neuron as the number of the dependent variables. But determining the neurons and structure of the hidden layer is difficult [6], [25].

Transferring function, which has been used for prediction by the most researchers, is: sigmoid, tangent hyperbolic and pureline [26], [27]. The researchers usually have applied the same transfer function in a layer, and occasionally have employed the same transfer function for all layers. In prediction problems, researchers usually have adopted Logistic function in the hidden layer and the linear function in the exit layer (see [27], [28]).

In this study, one hidden layer was employed and the number of neurons was determined by trial and error from one neuron to six neurons. The transfer function for neurons in the hidden layer was hyperbolic tangent. And the transfer function for the output layer was pureline. The structure which determines the least mean square error for the train sample was used as the final model. Nondiscretionary accruals were calculated by rendering the independent variables from a vector of the data to the input layer of the multi-layer perceptron. Then the discretionary accruals were calculated by distracting non-discretionary accruals from total accruals.

The number of neurons in the hidden layer was changed for the accuracy of the model. The classification of the data into three groups of training, validation, and test were 70%, 15%, and 15% respectively. Table III shows the result for every year. By considering the preceding approach, 54 models were designed and executed.

The schematic view of the MLP model in this research is shown in Fig. 1.

2. Generalized Regression Neural Networks (GRNN)

Similar to the Multi-Layer Perceptron (MLP), GRNN adopted a supervised learning algorithm. It consisted of one input layer, one hidden layer, one summation layer and one output layer. The number of neurons in input layer and exit layer were the same as the numbers of the independent and dependent variables. The train process of the GRNN was performed in a sweep. When some new data were entered in a GRNN, the distance between the input and weight vectors was calculated. Then this distance got through a radial function though a lower distance which made a greater exit value [6].

Dai et al. mentioned that the GRNN approach does not need any assumptions about the forms of the data [29]. So, the structure of the GRNN can be shown in Fig. 2.

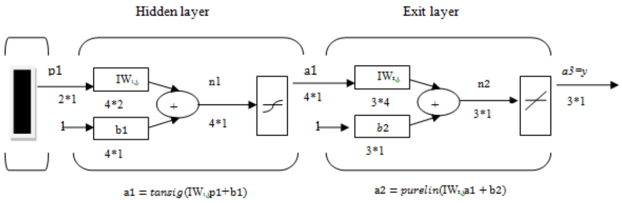


Fig. 1 The Structure of the MLP [31]

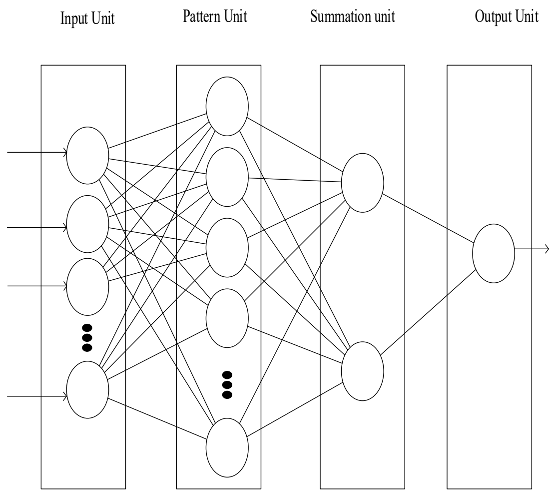


Fig. 2 The Structure of GRNN [29]

In applying neural networks techniques, however, the following factors should be considered:

1) When using neural networks, there is no clear rule for defining the structure and parameters of the networks, and the common way of defining the structure of the neural networks is based upon a trial and error [30],
2) Normally the train process of the neural networks is more time consuming than the regression models, and
3) Interpretation of the neural networks is different from regression models.

*C. Models' Comparison*

One problem in evaluating earnings management models, which are based on the accruals, is the difficulty of measuring the performance of the models. The reason behind this is that the true amount of earnings management is unknown. Hence, in this study, the performance of the models was evaluated based on the following criteria:

The discretionary accrual was estimated for the sample. It is expected that the mean and the median of the discretionary accruals for a large sample be near to zero. Therefore, the performance of these models was evaluated based on how these models would measure the mean of the discretionary accruals near to zero.

## V. Findings

### A. Descriptive Statistics

Table I shows the descriptive statistics of the study. It discloses that the maximum amount of the variables in most years (2003, 2004, 2005, 2007, 2008, 2010, and 2011), the minimum amount of the most years (2004, 2007, 2008, 2009, and 2010), and the maximum amount of the standard deviation for all years (except 2006) belongs to" changes in revenue divided by the last year total assets". This finding reveals the importance of the changes in revenue for examining earnings management.

TABLE I
DESCRIPTIVE STATISTICS OF RESEARCH VARIABLES

| year | variable | observation | Mean | maximum | minimum | Standard deviation |
|---|---|---|---|---|---|---|
| | 1/TAt-1 | 94 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| | Revenue t /TAt-1Δ | 94 | 0.2894 | 7.9464 | -0.2307 | 0.8451 |
| 2003 | PPE t /TAt-1 | 94 | 0.2941 | 0.9598 | 0.0004 | 0.2031 |
| | CFO t /TAt-1 | 94 | 0.1825 | 0.9252 | -0.2307 | 0.1909 |
| | TACC t /TAt-1 | 94 | 0.0759 | 0.3605 | -0.5104 | 0.2215 |
| | 1/TAt-1 | 93 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| | Revenue t /TAt-1Δ | 93 | 0.1599 | 1.1243 | -4.7065 | 0.5520 |
| 2004 | PPE t /TAt-1 | 93 | 0.2948 | 0.9084 | 0.0003 | 0.1937 |
| | CFO t /TAt-1 | 93 | 0.1889 | 1.0050 | 0.2405 | 0.2162 |
| | TACC t /TAt-1 | 93 | 0.0563 | 0.4561 | -0.4796 | 0.1550 |
| | 1/TAt-1 | 93 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | Revenue t /TAt-1Δ | 93 | 0.1929 | 2.0365 | -0.2259 | 0.3008 |
| 2005 | PPE t /TAt-1 | 93 | 0.2865 | 0.7728 | 0.0001 | 0.2010 |
| | CFO t /TAt-1 | 93 | 0.1156 | 0.6568 | -0.6151 | 0.1897 |
| | TACC t /TAt-1 | 93 | 0.1002 | 0.9071 | -0.3626 | 0.1764 |
| | 1/TAt-1 | 93 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | Revenue t /TAt-1Δ | 93 | 0.1195 | 0.8453 | -0.3828 | 0.2202 |
| 2006 | PPE t /TAt-1 | 93 | 0.2669 | 1.2228 | 0.0028 | 0.2065 |
| | CFO t /TAt-1 | 93 | 0.1443 | 1.3454 | -0.5322 | 0.2292 |
| | TACC t /TAt-1 | 93 | 0.0565 | 0.7364 | -0.5743 | 0.1812 |
| | 1/TAt-1 | 92 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| | Revenue t /TAt-1Δ | 92 | 0.1593 | 1.6376 | -0.8771 | 0.2951 |
| 2007 | PPE t /TAt-1 | 92 | 0.2616 | 0.8591 | 0.0052 | 0.1983 |
| | CFO t /TAt-1 | 92 | 0.1436 | 0.7143 | -0.7342 | 0.1989 |
| | TACC t /TAt-1 | 92 | 0.0622 | 1.1354 | -0.2024 | 0.1861 |
| | 1/TAt-1 | 93 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| | Revenue t /TAt-1Δ | 93 | 0.7637 | 55.4956 | -0.5105 | 5.7424 |
| 2008 | PPE t /TAt-1 | 93 | 0.2667 | 1.1460 | 0.0042 | 0.2187 |
| | CFO t /TAt-1 | 93 | 0.1454 | 0.7814 | -0.2333 | 0.1600 |
| | TACC t /TAt-1 | 93 | 0.2960 | 23.0929 | -0.3290 | 2.3939 |
| | 1/TAt-1 | 94 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | Revenue t /TAt-1Δ | 94 | -0.4142 | 0.5542 | -46.0206 | 4.7586 |
| 2009 | PPE t /TAt-1 | 94 | 0.2507 | 1.3358 | 0.0031 | 0.2234 |
| | CFO t /TAt-1 | 94 | 0.1315 | 0.7725 | -0.2346 | 0.1489 |
| | TACC t /TAt-1 | 94 | 0.02319 | 0.2779 | -0.2355 | 0.1093 |
| | 1/TAt-1 | 94 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | Revenue t /TAt-1Δ | 94 | 0.0607 | 1.2708 | -0.7719 | 0.2656 |
| 2010 | PPE t /TAt-1 | 94 | 0.2361 | 1.1094 | 0.0000 | 0.2000 |
| | CFO t /TAt-1 | 94 | 0.1163 | 0.6924 | -0.2655 | 0.1405 |
| | TACC t /TAt-1 | 94 | 0.0230 | 0.4635 | -0.2582 | 0.1138 |
| | 1/TAt-1 | 94 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | Revenue t /TAt-1Δ | 94 | 0.1226 | 1.8376 | -0.5234 | 0.3042 |
| 2011 | PPE t /TAt-1 | 94 | 0.2394 | 0.9420 | 0.0055 | 0.1872 |
| | CFO t /TAt-1 | 94 | 0.1069 | 0.9942 | -0.6996 | 0.1759 |
| | TACC t /TAt-1 | 94 | 0.0332 | 1.2335 | -0.8009 | 0.1923 |

TABLE II
THE SUMMARY INFORMATION OF MULTI-VARIABLE REGRESSION

| year | R | $R^2$ | Adjusted $R^2$ | Standard Error | Durbin-Watson |
|---|---|---|---|---|---|
| 2003 | 0.806 | 0.649 | 0.633 | 0.1341 | 1.996 |
| 2004 | 0.645 | 0.416 | 0.389 | 0.1211 | 1.696 |
| 2005 | 0.703 | 0.494 | 0.471 | 0.1283 | 2.163 |
| 2006 | 0.745 | 0.555 | 0.535 | 0.1236 | 2.161 |
| 2007 | 0.578 | 0.334 | 0.303 | 0.1553 | 1.909 |
| 2008 | 0.999 | 0.998 | 0.997 | 0.1212 | 1.903 |
| 2009 | 0.412 | 0.170 | 0.133 | 0.1018 | 2.241 |
| 2010 | 0.644 | 0.415 | 0.389 | 0.0889 | 1.765 |
| 2011 | 0.799 | 0.638 | 0.621 | 0.1183 | 2.021 |

TABLE III
THE MEAN SQUARE ERROR FOR MLP

| year | Number of neurons | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 2003 | Performance of train data | 0.0104 | 0.0222 | 0.0156 | 0.0127 | 0.0156 |
| | Performance of validation data | 0.0228 | 0.0399 | 0.0152 | 0.029 | 0.0119 |
| | Performance of test data | 0.0322 | 0.0211 | 0.0193 | 0.0122 | 0.0428 |
| | Total Performance of model | 0.0155 | 0.0247 | 0.0161 | 0.0315 | 0.0191 |
| 2004 | Performance of train data | 0.0094 | 0.0169 | 0.0082 | 0.0129 | 0.0092 |
| | Performance of validation data | 0.0145 | 0.0058 | 0.0097 | 0.0183 | 0.0095 |
| | Performance of test data | 0.0273 | 0.0158 | 0.0242 | 0.0165 | 0.0112 |
| | Total Performance of model | 0.0129 | 0.0151 | 0.0108 | 0.0142 | 0.0095 |
| 2005 | Performance of train data | 0.015 | 0.0093 | 0.0114 | 0.0101 | 0.0361 |
| | Performance of validation data | 0.0076 | 0.0066 | 0.0226 | 0.0086 | 0.0307 |
| | Performance of test data | 0.0084 | 0.0184 | 0.0168 | 0.0324 | 0.0244 |
| | Total Performance of model | 0.0129 | 0.0103 | 0.0139 | 0.0132 | 0.0335 |
| 2006 | Performance of train data | 0.0137 | 0.0089 | 0.0097 | 0.0239 | 0.0101 |
| | Performance of validation data | 0.0152 | 0.0304 | 0.0102 | 0.0178 | 0.0146 |
| | Performance of test data | 0.0089 | 0.0225 | 0.0273 | 0.0212 | 0.0284 |
| | Total Performance of model | 0.0132 | 0.0142 | 0.0124 | 0.0226 | 0.0135 |
| 2007 | Performance of train data | 0.0197 | 0.0074 | 0.0089 | 0.0191 | 0.0055 |
| | Performance of validation data | 0.0104 | 0.0176 | 0.0086 | 0.0167 | 0.015 |
| | Performance of test data | 0.0093 | 0.0532 | 0.0352 | 0.0562 | 0.0111 |
| | Total Performance of model | 0.0167 | 0.0159 | 0.0129 | 0.0244 | 0.0078 |
| 2008 | Performance of train data | 0.0093 | 0.0101 | 0.0108 | 0.0078 | 0.0076 |
| | Performance of validation data | 0.0175 | 0.0265 | 0.0086 | 0.0232 | 0.02 |
| | Performance of test data | 0.0203 | 0.0065 | 0.0193 | 0.0197 | 0.0103 |
| | Total Performance of model | 0.0122 | 0.012 | 0.0117 | 0.0119 | 0.0099 |
| 2009 | Performance of train data | 0.011 | 0.0103 | 0.0103 | 0.0077 | 0.0102 |
| | Performance of validation data | 0.0071 | 0.0169 | 0.0058 | 0.0092 | 0.0143 |
| | Performance of test data | 0.0061 | 0.0134 | 0.0102 | 0.0127 | 0.0147 |
| | Total Performance of model | 0.0097 | 0.0117 | 0.0097 | 0.0086 | 0.0115 |
| 2010 | Performance of train data | 0.0056 | 0.0068 | 0.0049 | 0.0061 | 0.0067 |
| | Performance of validation data | 0.144 | 0.007 | 0.0066 | 0.0045 | 0.0101 |
| | Performance of test data | 0.0138 | 0.0086 | 0.0107 | 0.0133 | 0.0055 |
| | Total Performance of model | 0.0081 | 0.0071 | 0.006 | 0.0069 | 0.007 |
| 2011 | Performance of train data | 0.0213 | 0.0074 | 0.0078 | 0.0151 | 0.0087 |
| | Performance of validation data | 0.0049 | 0.0229 | 0.0057 | 0.0111 | 0.0531 |
| | Performance of test data | 0.0314 | 0.0058 | 0.0274 | 0.0321 | 0.0133 |
| | Total Performance of model | 0.0203 | 0.0095 | 0.0104 | 0.017 | 0.016 |

TABLE IV
THE RESULT OF GRNN MODEL

| | | year | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
| spread | 0.1 | 0.0046 | 00049 | 0.0073 | 0.0051 | 0.0059 | 0.0061 | 0.0056 | 0.0033 | 0.0054 |
| | 0.2 | 0.0143 | 0.0107 | 0.012 | 0.0124 | 0.0102 | 0.0122 | 0.0091 | 0.0065 | 0.01 |
| | 0.3 | 0.0207 | 0.0146 | 0.0176 | 0.0179 | 0.0125 | 0.0153 | 0.0102 | 0.0089 | 0.0126 |
| | 0.4 | 0.0241 | 0.0174 | 0.0233 | 0.0215 | 0.0156 | 0.0171 | 0.0106 | 0.0101 | 0.0165 |
| | 0.5 | 0.0261 | 0.0192 | 0.0259 | 0.0241 | 0.0206 | 0.0181 | 0.0109 | 0.0108 | 0.0219 |
| | 0.6 | 0.0273 | 0.0203 | 0.0272 | 0.0266 | 0.0262 | 0.0188 | 0.011 | 0.0113 | 0.0281 |
| | 0.7 | 0.0281 | 0.0211 | 0.0281 | 0.0283 | 0.0294 | 0.0192 | 0.0111 | 0.0117 | 0.0314 |
| | 0.8 | 0.0287 | 0.0216 | 0.0286 | 0.0293 | 0.031 | 0.0195 | 0.0112 | 0.0119 | 0.033 |
| | 0.9 | 0.0291 | 0.022 | 0.0290 | 0.0300 | 0.032 | 0.0197 | 0.0113 | 0.0121 | 0.0339 |
| | 1 | 0.0294 | 0.0222 | 0.0293 | 0.0304 | 0.0326 | 0.0198 | 0.0113 | 0.0122 | 0.0344 |

### B. Inferential Statistics

Tables II-IV show the results of the selected models. Table II shows the summary information of the LR model for the designated years. Adjusted $R^2$ varies between 0.133 and 0.997. The maximum Adjusted $R^2$ belongs to the year 2008, and the minimum Adjusted $R^2$ belongs to the year 2009. As Table II shows, the Durbin Watson of the investigated variables is near to 2, which indicates that serial correlation problem does not exist among the designated variables.

Table III shows the MLP results for each year. The data is divided into three groups: train data, validation data, and test data. The percentages were 70%, 15%, and 15% respectively. The number of neurons in the hidden layer was changed from two to six neurons. As Table III shows, the best performance for 2003 occurred in a model with five neurons with the mean square error of 0.01225. In year 2004, the best performance occurred in the model with six neurons with the mean square error of 0.0112. In the years from 2005 to 2007, the best performance occurred in models with two neurons with the mean square error of 0.0084, 0.0089, and 0.0093 respectively. The best performance of the year 2008 occurred in a model with three neurons with the mean square error of 0.0065. In the year 2009, the model with two neurons had the best performance with a mean square error of 0.0061. In year 2010, the best performance occurred in a model with a six neurons and the mean square error of 0.0055. The model with a three neurons and the mean square error of 0.0058, had the best performance in the year 2011

Finally, the result of the GRNN is shown in Table IV. The spread changed among 0.1 to 1 with the interval of 0.1. Therefore, 90 models were built to obtain the results. As Table IV indicates, the best model for all years is where the spread is 0.1.

After obtaining related data for the designated models, research hypotheses were investigated. The less the mean square error, the more precise is the model. Table V shows the result of the different models for detecting earnings management.

TABLE V
THE MEAN SQUARE ERROR FOR THE RESEARCH MODELS

| year | Linear regression | MLP | GRNN |
|---|---|---|---|
| 2003 | 0.018 | 0.01225 | 0.0046 |
| 2004 | 0.015 | 0.0112 | 0.0049 |
| 2005 | 0.016 | 0.0084 | 0.0073 |
| 2006 | 0.015 | 0.0089 | 0.0051 |
| 2007 | 0.024 | 0.0093 | 0.0059 |
| 2008 | 0.015 | 0.0065 | 0.0061 |
| 2009 | 0.010 | 0.0061 | 0.0056 |
| 2010 | 0.008 | 0.0055 | 0.0033 |
| 2011 | 0.014 | 0.0058 | 0.0054 |
| mean | 0.0150 | 0.00822 | 0.0053 |

At first, the normality of the mean square error was investigated by conducting Kolmogrov-Smirnov test and Shapiro-Wilk test. Table VI posits the results.

Since the significance level of every model was above the 0.05, the normality of the data was finalized.

### C. Hypotheses Testing

Consequently, for accepting or rejecting the hypotheses, the one-way ANOVA was exerted. Table VII shows the result of the first hypothesis. It indicates that the difference between MLP and GRNN were insignificant, and therefore H1 was not accepted. Hence, the precision of the GRNN was not greater than the precision of the MLP in predicting earnings management.

Table VIII shows the result of the second hypothesis. It indicates that the difference between MLP and LR model is significant and therefore, H2 was accepted. In other words, the precision of the MLP was greater than the precision of the LR model in predicting earnings management

Table IX shows the result of the third hypothesis. It indicates that the difference between GRRN and LR model is significant. By considering the mean square error of the GRNN (0.00536), which was less than the mean square error of the linear regression model (0.0150), H3 was accepted. Hence, the precision of GRNN was greater than the precision of the LR model in predicting earnings management.

TABLE VI
THE NORMALITY TEST

| Model | Kolmogrov -Smirnov Test | | | Shapiro-Wilk Test | | |
|---|---|---|---|---|---|---|
| | statistic | Degree of freedom | sig | statistic | Degree of freedom | sig |
| Linear Regression | 0.191 | 9 | 0.200 | 0.928 | 9 | 0.459 |
| MLP | 0.204 | 9 | 0.200 | 0.914 | 9 | 0.342 |
| GRNN | 0.139 | 9 | 0.200 | 0.976 | 9 | 0.941 |

TABLE VII
ANOVA RESULT FOR GRNN AND MLP

| Variable 1 | Variable 2 | Variable 1- Variable 2 | Standard Error | Sig |
|---|---|---|---|---|
| GRNN | MLP | -0.00286 | 0.00330 | 0.391 |

TABLE VIII
ANOVA RESULT FOR MLP AND LINEAR REGRESSION

| Variable 1 | Variable 2 | Variable 1- variable 2 | Standard error | sig |
|---|---|---|---|---|
| MLP | Linear Regression | -0.00678 | 0.00330 | 0.047 |

TABLE IX
ANOVA RESULT FOR GRNN AND LINEAR REGRESSION

| Variable 1 | Variable 2 | Variable 1- variable 2 | Standard error | sig |
|---|---|---|---|---|
| GRNN | Linear Regression | -0.00964 | 0.00330 | 0.006 |

## VI. CONCLUSION AND DISCUSSION

The major purpose of this study was comparing the precision of the Linear Regression (LR) with Artificial Neural Network (ANN) models in predicting earnings management in Tehran Stock Exchange (TSE) Market.

The result of investigating the first hypothesis showed that there was not a significant difference between the precision of the Generalized Regression Neural Networks (GRNN) and Multi-Layer Perceptron Neural Networks (MLP) in predicting earnings management. This finding is the opposite of the Höglund's conclusions [6]. A major reason for this opposite finding might be due to the number of observations, public nature of the firms listed in TSE, and the efficiency of the TSE market.

The result of investigating the second hypothesis showed a superior performance of the MLP in contrast with the LR model. This superiority may lie on the nonlinear relation which exists between the independent and dependent variables. The acquired results are in accordance with findings of [6], and [32].

The research results also showed the acceptance of the third hypothesis. Therefore, a superior performance of the GRNN over the LR model was shown. The reason might also be due to the nonlinearity relationship between the independent and dependent variables. This result is consistent with the Höglund research [6].

In sum, this research showed that neural networks models (GRNN and MLP) could more accurately predict earnings management than the linear regression model (LR) model. This finding is the opposite of [17]-[19]. Its implication is very important since it provides an empirical evidence concerning the usefulness of the neural network models in a developing (TSE) market. It also indicates that earnings management could be analyzed more accurately by considering nonlinear relationships among the designated variables. In addition, it provides a useful tool and means for the investors, financial analysts and capital markets participants, and helps them to predict earnings management more accurately. Since the greatest loss to the capital market, up to the present time, has been as a result of the earnings management, it would prevent the occurrence of the unfavorable effects of earnings management. By applying neural networks models, market participants can predict the earnings management, and are not mislead by the opportunistic behavior of the managers. On the other hand, managers who do follow the opportunistic behavior will, sooner or later, find out that capital market participants will know the fact, and therefore will become more conservative to implement earnings management. This will ultimately lead to a lower rate of exercising earnings management by management.

## VII. SUGGESTIONS

Based upon the results of the study, the following suggestions are made:

1) Since the best results are acquired from adopting the GRNN and MLP, it is suggested that researchers apply these models of neural networks to obtain more accurate prediction results.
2) Tehran Stock Exchange companies and financial analysts should use neural networks techniques for detecting earnings management.
3) Due to the importance of the earnings management and also improving the methods of detecting earnings management, the following suggestions are made:

a) Investigating the current study in a broader period of time.
b) Detecting earnings management by using other techniques of artificial intelligence.
c) Duplicating the current study with different learning functions and also different structures of the neural networks.

## REFERENCES

[1] Vladu A. B. and D. D. Cuzdrioreanb (2014). “Detecting earnings management: Insights from the last decade leading journals published research” Procedia Economics and Finance, Vol. 15, pp. 695-703.
[2] Graham, J. R., Raedy J. S., and D. A. Shackelford, (2012). “Research in accounting for income taxes”, Journal of Accounting and Economics, Vol. 53, pp. 412-434.
[3] Jansen, I. P.; Ramnath, S. and T. L. Yohn (2012) “A Diagnostic for Earnings Management Using Changes in Asset Turnover and Profit Margin.” Contemporary Accounting Research, Vol. 29, No. 1, pp. 221–251.
[4] Capkun, V. (2011). “Book Review.” The International Journal of Accounting, Vol. 46, No. 2, pp. 236-237.
[5] Wróblewski D., Jarne, José I. and S. Callao (2014). “The Development of Earnings Management Research: A Review of Literature from Three Different Perspectives” Theoretical Journal of Accounting (Zeszyty Teoretyczne Rachunkowości), Vol. 79, pp. 135-177.
[6] Höglund, H. (2012). “Detecting Earnings Management with Neural Networks.” Expert Systems with Applications, Vol. 39, pp. 9564-9570.
[7] Höglund, H. (2013). “Fuzzy Linear Regression-Based Detection of Earnings Management.” Expert Systems with Applications, Vol. 40, pp. 6166-6172.
[8] DeTienne, K. B., DeTienne, D. H., and S. A. Joshi (2003). “Neural Networks as Statistical Tools for Business Researchers” Organizational Research Methods, Vol. 6(2), pp. 236–265.
[9] Tsai C. F. and Y. J. Chiou (2009). “Earnings Management Prediction: A Pilot Study of Combining Neural Networks and Decision Trees.” Expert Systems with Applications, Vol. 36, pp. 7183–7191.
[10] Dastgir, M. and A. Nazemi (2006). “Earnings Management: Reconciling the Views of Accounting Academics, Practitioners, and Regulators: Review of Iranian Studies” Journal Of Accounting Knowledge and Study. Vol. 11, pp. 12-18.
[11] Ronen, J. and V. Yaari (2008) Earnings Management: Emerging Insights in Theory, Practice, and Research, Springer, the USA.
[12] Scott, William R. (2009). Financial Accounting Theory. Fifth Edition. Upper Saddle River, NJ: Prentice Hall.
[13] Hribar, P., and D. W. Collins (2002). "Errors in Estimating Accruals: Implications for Empirical Research." Journal of Accounting research 40(1), pp. 105-134.
[14] Healy, P. M. and J. M. Wahlen (1999). “A Review of the Earning Management Literature and Its Implications for Standard Setting.” Accounting Horizons, Vol. 13, PP. 365-383.

[15] Bergstresser, D. and T. Philipponb (2006). “CEO Incentives and Earnings Management.” Journal of Financial Economics, Vol. 80, pp. 511–529.
[16] Namazi M. and E. Khansalar (2011). “An Investigation of The Income Smoothing Behavior of Growth and Value Firms (Case Study: Tehran Stock Exchange Market).” International Business Research, Vol. 4, No. 4, pp. 84-93.
[17] Denton, J. W. (1995). “How Good are Neural Networks for Causal Forecasting?” Journal of Business Forecasting Methods and Systems, Vol. 14(2), pp. 17–20.
[18] Warner, B.; and M. Misra (1996). “Understanding Neural Networks as Statistical Tools” The American Statistician, Vol. 50(4), pp. 284–293.
[19] Marquez, L., Hill, T., Worthley, R., and W. Remus, (1992). Neural network models as an alternative to regression.In R. Trippe & E. Turban (Eds.), Neural Networks in Finance and investing (pp. 435–449). Chicago: Probus.
[20] Namazi M, Shokrolahi A, and M. Sadeghzadeh Maharluie. (2014). "Ranking Relative Importance of Various Determinants of the Free Cash Flow Risk: A Neural Network Approach" the 4th African Accounting & Finance Association Conference, South Africa, Cape Town, September 3-5.
[21] Rees, Lynn L., Gill, Susan and Richard Gore (1996). “An Investigation of Asset Write-Downs and Concurrent Abnormal Accruals” Journal of Accounting Research, Vol. 34, pp. 157-169.
[22] Ball R. and L. Shivakumar (2006). “The Role of Accruals in Asymmetrically Timely Gain and Loss Recognition.” Journal of Accounting Research, Vol. 44, No. 2, pp. 207-242.
[23] Aaker, H. and F. Gjesdal (2007). “Detecting True Earnings Management and Evaluating Discretionary Accrual Models” International Conference, Handelshögskolan vid Göreborgs University.
[24] Kouki, M.; Elkhaldi A.; Atri, Hanen; and S. Souid (2011). “Does Corporate Governance Constrain Earnings Management? Evidence from U.S. Firms” European Journal of Economics, Finance and Administrative Sciences, Vol. 35, pp. 58-71.
[25] Menhaj, M. B. (2010). Introductory to Neural Networks, Seventh Edition, Tehran, Amirkabir University of Technology.
[26] Duch, Wlodzislaw and Norbert Jankowski (1999). “Survey of Neural Transfer Function”, Neural Computing Surveys, Vol. 2, pp. 163-212.
[27] Namazi, M. and M. M. Kiamehr (2007). “Predicting Daily Stock Returns of Companies Listed in Tehran Stock Exchange Using Artificial Neural Networks”. Journal of Financial Research, Tehran University, Vol. 24, pp. 115-134.
[28] Moshiri, S. and H. Morovat (2005). “Predicting Tehran Stock Returns Index by using Linear and Nonlinear Models” Trade Studies Quarterly, Vol. 41, pp. 245-276.
[29] Dai, J.; Liu, X; Zhang, S.; Zhang, H.; Xu, Q.; Chen, W. and X. Zheng (2010). “Continuous Neural Decoding Method Based on General Regression Neural Network” International Journal of Digital Content Technology and its Applications, Volume 4, Number 8, pp. 216-221.
[30] Zhang, G., Patuwo, B. E., and M. Y. Hu, (1998). “Forecasting with Artificial Neural Networks: The State of the Art” International Journal of Forecasting, Vol. 14(1), pp. 35–62.
[31] Hudson Beale, Mark; Hagan Martin T. and Howard B. Demuth (2012). Neural Network Toolbox™ User’s Guide, The MathWorks, Inc.
[32] Hejazi, R; Mohammadi, S; Aslani, Z. and M. Aghajani (2012). “Earnings Management Prediction Using Neural Networks and Decision Tree in TSE”. Iranian Accounting and Auditing Review, Vol. 19 (2), pp. 31-46.