

Application of Artificial Intelligence Model to Identify the Distorted Financial Statements

Behzad Soleymanian Asl¹, Razieh Solgi^{2,3*}

¹Department of accounting, Attar University of Mashhad, Mashhad, Iran

²Department of Medical Physics and Biomedical Engineering, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran

³Research center for molecular and cultural imaging, Tehran University of Medical Sciences, Tehran, Iran

neginakhondi103@gmail.com

Abstract

Distortion of financial statements is recognized as one of the most important issues in the field of accounting and auditing, which is also one of the most common issues today. In this regard, the present research was conducted, in which stock exchange information was used to investigate, predict, and model accounting distortions. For this purpose, financial performance, non-financial metrics, market-based metrics and commitment, or selection items were reviewed over a 6-year period. For collecting data of distorting companies, database of the Society of Certified Public Accountants in Iran was used and the information was analyzed using data mining methods (decision tree, neural networks, and Bayesian method). The results showed that analysis of financial statements' information has a high accuracy in determining and identifying the distorted financial statements. Using this information, it is possible to get better acquainted with the methods of document distortion and to take necessary measures in order to control and prevent administrative violations at national and international levels. Given frequent occurrence of these violations, artificial intelligence models can be used to identify these papers.

Keywords: Accounting Distortions, Financial Statements, Neural Network, and Accounting.

Introduction

The detrimental effects of financial scandals in the recent years have increased attention towards the issue of fraud and distortion in financial statements. Accounting misconduct is referred to as any discrepancy between amount, classification, presentation, or disclosure of an item reported in financial statements and amount, classification, presentation, or disclosure of that item in accordance with requirements of the accepted financial reporting framework (international federation of accountants (IFAC), December, 2009: 370). Distortions are the result of mistakes or fraud and a mistake is defined as an inadvertent act in financial statements including omission of an amount or a case of disclosure and fraud is defined as any intentional or fraudulent act by one or more directors or third parties to obtain an unlawful advantage, according to the 2008 report of association of certified fraud examiners (ACFE). It is estimated that fraud schemes involve losses costing \$ 1 billion a year. In

addition to economic losses, such as harming creditors, investors, and shareholders, financial scandals have caused political, judicial, and social costs and due to the decrease in reliability of companies' financial statements, it has led to the increased transfer costs and the reduced efficiency of capital market (Prolls, 2011). Rezaei and Riley (2010) argued that extent of fraudulent financial statements has raised public concerns and undermined public trust in financial reporting process and auditing performance. This phenomenon threatens quality, integrity, and reliability of financial reporting process and causes significant economic losses to investors and creditors (Sajjadi and Kazemi, 2016). According to the 2016 report by the ACFE, up to 23 % of the surveyed subjects had lost more than \$ 1 million of their capital. Misappropriation of asset is the most common type of fraud, accounting for 84% of total fraud cases on average and causing \$ 125,000 losses on average to each economic unit. Fiscal statements, on the other hand, account for only 10 % of all fraud cases, causing \$ 975,000 losses per unit on average. Because of importance of this issue, extensive research has been conducted to understand causes, motives, and consequences of financial distortion and profit manipulation (Kim et al., 2016). In particular, determining how financial statements and income are detected is a major concern for research on the field of accounting. Therefore, results of this study, in addition to expanding the research literature on the distorted financial statements can increase confidence of users of financial statements and thus, increase operational efficiency of capital market. In other words, protecting interests of business units stakeholders and helping to improve internal mechanisms of corporate governance as well as securities markets regulators in identifying and predicting the possibility of distortion in financial statements before its occurrence are among reasons for importance of conducting the research. For improving prediction of distortion, research in the field of accounting has focused on application of various statistical methods. Thus, distinguishing features of this research compared to the previous research are as follows: First; in this study, an attempt was made to investigate a large number of financial and non-financial variables studied in the previous research simultaneously. Second; in the previous research, mainly due to lack of publishing names of companies distorting financial statements, the distorted companies have been identified using other criteria, but in this research, database of the Society of Certified Public Accountants was used. Third; in most studies, regression method has been used to predict distortion and other data mining methods, such as decision tree, neural networks, and Bayesian method have not been used frequently. Therefore, this study is conducted to investigate the possibility of predicting accounting distortions using decision tree methods, neural networks, and Bayesian method. To this end, companies with accounting distortions are predicted and identified by analyzing characteristics of companies that have been distorted in the past - according to the four dimensions of accruals, financial performance, non-financial criteria, and market-based criteria - and reviewing the research conducted by Decho et al., (2011). Variables are classified and tested in three different models. The first model includes variables of accruals and company performance. In the second model, in addition to accruals and company performance; non-financial variables are also considered. In the third model, in addition to the previous variables, market-related variables are added and the following question is addressed: Do the decision tree, neural networks, and Bayesian methods allow predicting the distorted financial statements. Many authors use the term "distortion" for manipulation of financial statements, resulting in unfair presentation of

financial position of the business. The official definition of accounting distortion has been provided by IFAC as follows: Any discrepancy between amount, method of classification, presentation or disclosure of an item reported in financial statements and amount, classification, presentation or disclosure of that item in accordance with requirements of the accepted financial reporting framework (December, 2009,: 370). This definition is also in accordance with Iran's Auditing Standard No. 240. Distortions are divided into intentional and unintentional distortions. In other words, distortions are the result of a mistake or fraud, which is an inadvertent act in financial statements including omission of an amount or a case of disclosure and fraud includes any intentional or fraudulent attempt by one or more directors or third parties to obtain an unlawful advantage. Most models of detecting financial distortions face with a dilemma. As can be seen in Fig. 1, the companies with financial distortion (or subset A as fraudulent companies) and the companies that have not been distorted (i.e., A vs. D or classification of companies into two categories of the distorted companies and not-distorted companies (C vs. D where C includes A and B) are presented in this study (Kim et al., 2016).

Reviewing the research literature, it can be concluded that many studies have been done using descriptive and forecasting techniques to predict simultaneous accounting distortions. In this research, prediction techniques were used to detect distortion.

Materials and Methods

This research is applied in terms of purpose and of experimental and post-event type, which is based on real information of financial statements of companies listed in the Tehran Stock Exchange and data mining methods including decision tree, neural networks, and Bayesian learning were selected to analyze the results.

Decision Tree

Learning decision tree is one of the most common deductive inference algorithms that has been successfully utilized in a wide range of applications, from medical diagnoses to assessment of credit risk. The decision tree, whose main purpose is categorizing data, is a model in data mining that provides a tree-like structure for making decisions and determining class of a particular data, like a flowchart. As its name suggests, this tree is composed of a number of nodes and branches (Qaderzadeh et al., 2017). The leaves represent categories, and middle nodes are used to make decisions about one or more specific attributes. An important advantage of the decision tree algorithm is its easy interpretation and comprehensibility. If the tree is allowed to grow indefinitely, in addition to spending a lot of time, it will cause the tree to over-adapt to training data and therefore, the developed decision tree will not be generalizable. Size of the trees can be controlled by stop rules or the tree can be pruned after construction. REPTree algorithm is one of well-known decision tree learning algorithms herein, its good results will be presented for modeling problems' data. REPTree algorithm is a fast classification algorithm that is based on gaining information and calculation of entropy.

Neural Networks

Neural networks are one of the most widely used methods for modeling complex and large problems. A neural network is a data processing system that takes ideas from the human's brain and delegates data processing to many small processors acting as interconnected, parallel networks to solve a problem. In these networks, a data structure is designed with the help of programming knowledge that can act like a neuron, called as a node. In the neural network method, the assumption of non-linearity and independence of explanatory variables is eliminated, in which the hidden relationships between explanatory variables enter the function as an additional variable. Neural networks have different architectures. One of the most popular neural network architectures is multilayer perceptron because; it can consider both non-linearity and interaction of variables.

Bayesian Learning

In general, the purpose of Bayesian learning and deductive learning is finding the best hypothesis "h" in the H space using the "D" training set. The best hypothesis is the hypothesis that is most probable with respect to the D dataset. Given this definition, the problem of machine learning will become an attempt to determine the probability of different H hypotheses and to select the highest probability based on D dataset. Conditional probability, $P(h | D)$ is the probability of hypothesis "h" after observing D dataset. Bayes' theory, which is given in the following equation, intends to calculate this conditional probability and is the basis of Bayesian learning. The main advantage of this method is that it allows determining the probability for occurrence of a particular event based on a set of actions, and in this way, a clear view of the relationship will be obtained.

$$p(h|D) = \frac{p(D|h)p(h)}{p(D)}$$

Unbalanced Data

Unbalanced datasets are the data, in which distribution of samples between different classes is unbalanced. In other words, in this type of problem, the number of samples in certain categories is much less than other categories. For example, in the problem studied in this paper, the number of distortion class data is much less than normal class data. Machine learning algorithms usually work poorly on the unbalanced data because they can easily provide a model that predicts all samples belonging to a larger class. In other words, due to the lack of data of distorting companies, model output of non-distorting companies is well predicted and it has little ability to predict distorting companies, which is the main purpose of building the model. Therefore, data modification or learning method seems necessary for solving this type of issues.

The Dependent Variable

Distortion in financial statements is considered as dependent variable, which has a qualitative nature and has a nominal scale. For measuring this variable, the value of 1 is assigned to distorting companies and the value of 0 is assigned to other companies. In this research, the

database of the Society of Certified Public Accountants was used to identify and classify distorting economic units.

Independent Variables

Independent variables of the research are as follows:

1. Accruals

For the first time, Heli (1985) stated that profits are distorted through accruals. Therefore, in this study, the issue stating that whether the years, in which the profit is distorted are related to emergence of unusual high accruals or not is investigated. Based on the previous research, 12 variables have been used as a criterion for accruals including working capital accruals, Jones Model's Richardson, Sloan, Soliman, and Tuna (RSST) accruals, Modified Jones Model, Alar Kerr & Richardson Model, Kotari Model, Decho & Digo Model, McNichols Model, change in receivables, change in inventories, and percentage of soft assets and the first accrual measure focuses on working capital accruals.

1-1) Working Capital Accruals

$$WC-ACC = [\Delta CA - \Delta CSI] - [\Delta CL - \Delta DCL - \Delta TP] / ATA$$

WC - Acc: Working capital accruals; ACA: Changes in the current assets; ACSI: Cash changes and short-term investments; ACL: Current debt changes; DCL: Changes in the current liabilities; 2TP: Payable tax changes; and ATA: Average total assets.

1 - 2) RSST Accruals

The next benchmark, abbreviated as RSST, was introduced in 2005 by Richardson et al., and it was expanded into working capital accruals including changes in long-term operating assets and long-term operating liabilities. This scale is consistent with change in net non-cash assets.

$$rsst-acc = \frac{(\Delta WC + \Delta NCO + \Delta FIN)}{ATA}$$

1-3) Decho and Digo Model

$$\Delta WC_{it} = B_0 + B_1 CFO_{it-1} + B_2 CFO_{it} + B_3 CFO_{it+1} + \epsilon_{it}$$

AWC_{it} is the change in working capital in this year compared to the last year, 1-CFO_{it} is operating cash flow in the previous year, CFO_{it} is operational cash flow in this year, and CFO_{it} + 1 is operational cash flow in the next year.

1-4) McNichols Model

$$\Delta WC_{it} = B_0 + B_1 CFO_{it-1} + B_2 CFO_{it} + B_3 CFO_{it+1} + B_4 (\Delta REV_{it}/A_{it-1}) + B_5 (PPE_{it}/A_{it-1}) + \epsilon_{it}$$

1-5) Change in Receivables

It is referred to as change in accounts receivable and documents, accompanied by a distortion to show sales growth and is related to investors.

$$Ch-rec = \Delta AR / ATA$$

AAR: Change in accounts receivable and documents and ATA: Average total assets.

1-6) Change in Inventories

$$Ch-inv = \Delta INV / ATA$$

AINV: Change in inventories and ATA: Average total assets.

1-7) Percentage of Soft Assets

$$Soft-assets = (TA - PP\&E - CE) / TA$$

TA: Total assets; PP & E: Property, plant, and equipment; CE: Cash and cash equivalent, and TA: Total assets

2) Performance

Another set of variables is financial performance of the company, evaluated from different dimensions, and herein, the question of whether managers distort financial statements to cover up poor financial performance is investigated.

2-1) Change in Cash Sales

$$Ch-cs = SAAR$$

S: Sales and AR: Change in receivables.

2-2) Change in Cash Profit Margin

$$Ch-cm = 1 - [(CGS - AINV + AAP) / (S - AAR)]$$

CGS: Cost of goods sold; AINV: Change in inventories; AAP: Change in accounts payable and commercial documents; S: Sales and AAR: Changes in accounts receivable and business receivables.

2-3) Change in Rate of Return on Assets

$$CH-ROA = \left[\frac{E_T}{ATA_T} \right] - \left[\frac{E_{T-1}}{ATA_{T-1}} \right]$$

2-4) Change in Free Cash Flows

$Ch-fcf = A [E - RSST \text{ Accruals}] / ATA$, Ch - fcf: Change in free cash flows; E: Net profit; RSST Accruals: RSST accruals; and ATA: Average total assets.

3) Non-Financial Criteria

$CH-EMP = AEMPLOYEE - ATA$, Ch - emp: Abnormal change in the number of employees; Employee: Percentage of change in the number of employees, and ATA: Percentage of change in assets

4) Motivations Related to Markets

4-1) Publication of Shares

If the company has issued shares in the year in question, the value of 1 is assigned to it, otherwise the value of 0 is assigned to it.

4-2) Then, net amount of new financing is considered in relation to ATA.

$Cff = \text{Level of finance raised} / \text{ATA}$, in which amount of financing has increased.

4-3) Financing Needs (EXFIN)

$$\frac{(\text{CFO} - \text{past three years average capital expenditures})}{\text{current assets}} \leq -0.5$$

4 - 4) Financial Leverage

$$LEV = LD / TA$$

LD: Long-term debt and TA: Total assets.

5-4) Ratio of Book Value to Market Value

$$bm = \text{Equity}_t / MV_t$$

Equity: This year's equity and MV: This year's stock market's value.

6-4) Ratio of Profit to Price

$$ep = E / MV$$

In this regard, E is net profit and MV is stock market's value.

Population and Statistical Sample

Statistical population of this study consisted of all companies listed in the Tehran Stock Exchange whose information was available during the period of 2009-2015 and were not part of banks and financial institutions of investment companies, financial intermediation holding companies, and leasing companies. According to the above criteria, 189 companies including 21 distorting companies and 168 control companies were selected.

All companies	412
Insurance, investment, banking ,and financial intermediation companies	28
Companies whose end of fiscal year is not equivalent to the end of March	52
Companies whose information was not available	142
Number of the surveyed companies	189

Results

Table 1 presents descriptive statistics of the research variables including mean, maximum, minimum, and standard deviation for distorting and control companies. SPSS 23 software was used to compare the two groups. Kolmogorov-Smirnov test was used to check normality of data distribution. Due to non-normality of the research data, Mann-Whitney U test was used to compare the two groups of distorting and control companies. The results showed a statistically significant difference in average financial leverage between the two groups of companies at significance level of 5%. On average, larger financial leverage was reported in distorting companies. Jones et al., (2008) in a study showed that distorting firms have reported less financial leverage, which is probably due to differences in the way foreign companies are financed compared to the Iranian companies. The difference in average residual accruals in all accruals models homogenized by the first assets of the period was significant at significance level of 5%, which is consistent with the research by Jones et al., (2008).

Also, regarding other variables, no significant difference was observed between companies of the two groups, and the criteria of optional accruals calculated in all models were able to predict distortion, and the biggest difference between the distorting and control companies was related to the criteria of Decho and Digo, and McNichols Models (about 68%) and therefore, it seems that these two criteria better differentiate between accruals models in distorting companies.

The software was used for modeling as well as cost-sensitive learning to increase sensitivity of data mining methods used for distortion class. Pruning is also used in the REPTree algorithm to prevent over-compliance. For modeling neural networks, multilayer perceptron method with a latent layer was used. Also, the multilayer perceptron architecture was adjusted with one hidden layer and 4 hidden nodes. Also, learning and momentum rates were considered equal to 0.3 and 0.2, respectively. The number of repetitions in network learning process was set as 500 times. The simple Bayesian method was used to model the data under study.

Results of Testing the First Model

The cost of detecting normal distorting companies for modeling by neural network, Bayesian method, and decision tree in the first model was equal to 65, 50, and 70, respectively, obtained by trial and error. As shown in Table 2, average overall accuracy of the decision

tree, neural networks, and Bayesian method was equal to 54, 86, and 87%, respectively. Average correct detection rates were obtained as 48, 62, and 56%, respectively, and subsurface level of rock was measured as 52, 62%, and 55%, respectively. Therefore, the first model of neural network method had the highest ability to predict distortion (with an overall accuracy of 86 %, correct detection rate of 62 %, and area under the curve of 62 %).

Results of Testing the Second Model

Detection cost of normal distorting companies for modeling by decision tree, neural networks, and Bayesian method in the second model was equal to 60, 50 and 70, respectively. As can be seen in Table 3, the overall accuracy of the decision tree, neural networks, and Bayesian method is equal to 62, 80, and 87%, respectively, and underside of the rock is equal to, respectively:

Results of Testing the Third Model

Cost of detecting normal distorting companies for modeling by the neural network, Bayesian method, and decision tree in the third model was equal to 60, 45, and 70, respectively. As can be seen in Table 4, the overall accuracy of the decision tree, neural networks, and Bayesian method is 73, 76, and 86%, respectively, and the subsurface level of rock is equal to 50, 50, and 57%, respectively. Therefore, in the third model, the Bayesian method had the highest ability to predict distortion (with an overall accuracy of 86%, a subsurface level of 57%, and detection accuracy of 57%).

Discussion and Conclusion

Predicting and identifying companies distorting financial statements is one of the most important issues in the field of accounting. Very satisfactory results can be achieved by anticipating and preventing distortion. In this research, the decision tree model, neural networks, and Bayesian method, as data mining methods were used to predict distortion. The results of study showed that the information of financial statements has a high predictive power to predict distortion and therefore, it is possible to predict distortion only using the information of financial statements. In other words, using accruals and financial performance, the data of which can be extracted from financial statements, the distorted financial statements can be identified. Our results showed a relationship between all models of accruals and the distorted financial statements. The results of the study are consistent with the results of the research conducted by Mashayekhi and Hassanpour (2016) and Hestan et al., (2013) on the use of accruals by companies to distort financial statements. The results also indicated that the neural networks and Bayesian methods are more accurate than the decision tree. In other words, the Tehran Stock Exchange Organization or the Society of Certified Public Accountants can produce useful software to predict companies that may intend to distort financial statements and gain trust of shareholders by increasing their control over them. Shareholders and creditors, as individuals who are subject to irreparable damage by distortion can also reduce their losses in this way. Notably, these companies were removed in the selected sample due to nature of intermediary companies, making it impossible to generalize the research findings to all companies and the issue in such companies should be considered

separately. Lack of access to information of non-listed companies was another limitation of the research making it impossible to generalize results to all companies. Also, another limitation of the research was the lack of access to all distorting companies due to non-publication by auditing organization or lack of access to their information in the years when their distortion has been proven. In the future research, the results can be improved by adding corporate governance features. Also, considering other methods of detecting distortion, such as presenting more or less real profit and separating real profit management from accrual profit management can lead to the improved results in the future research. In addition to reviewing items of the financial statements distorted by managers, determining frequency of how to distort financial statements can also help to predict distortion. While collecting data from companies distorted financial statements, there was a problem regarding not providing information or not providing information in a timely manner, according to requirements of the stock exchange organization.

منابع

- پورحیدری، امید. سمیعی نژاد، نفیسه. (۱۳۹۴). بررسی تاثیر ساختار حاکمیت شرکتی بر گزارشگری مالی. پژوهش های کاربردی در گزارشگری مالی، ۶، ۵۳-۸۰. رحمانی، علی. اسماعیلی، غریبه. (۱۳۸۹). کارایی شبکه های عصبی رگرسیون لجستیک و تحلیل تمایزی در پیش بینی نکول. مجله اقتصاد مقداری، ۷ (۴).
- زجاجی، زهرا. (۱۳۸۷). تشخیص نفوذ در شبکه های کامپیوتری با استفاده از یادگیری حساس به هزینه. پایان نامه کارشناسی، دانشگاه اصفهان. سجادی، حسین کاظمی، توحید. (۱۳۹۵). الگوی جامع گزارشگری مالی متقلبانه در ایران
- به روش نظریه پردازی زمینه بنیان. پژوهش های تجربی حسابداری، ۲۱، ۲۰۴-۱۸۵. شمس الدینی، کاظم. دانشی، وحید. محلاتی راینی، محمد علی. (۱۳۹۷). توانایی قانون
- بنفورد در کشف تقلب در دو صورت سود یا زیان و صورت وضعیت مالی، مجله دانش حسابداری، ۷۳، ۱۶۷-۱۸۶. عبدالله زاده، سلام. محمدی ملقرنی، عطالله. نوروش، ایرج. امینی، پیمان. (۱۳۹۷). بررسی تاثیر آنتروپی نسبت های نقدینگی و اهرمی بر شناسایی شرکت های مشکوک به تقلب
- حسابرسی، ۷۲، ۷۹-۹۶. در گزارشگری مالی. مجله دانش قادرزاده، سید کریم. کردستانی، غلامرضا. حقیقت، حمید. (۱۳۹۶). شناسایی و رتبه بندی
- عوامل مؤثر بر سطح افشای مسئولیت اجتماعی شرکتها با رویکرد داده کاوی.
- پژوهش های کاربردی در گزارشگری مالی، ۱۱، ۷-۴۶. کردستانی، غلامرضا. تاتلی، رشید. (۱۳۹۵). پیش بینی دستکاری سود: توسعه یک مدل
- مجله بررسیهای حسابداری و حسابرسی، ۲۳، ۹۶-۷۳.

Abdollah Zade, S. Mohamadi Molgharni, A., Nouravesh, I., Amini, P. (2018) . Investigating the effect of entropy of liquidity and leverage ratios on identifying companies suspected of fraud in financial reporting. *Journal of Audit Science*, 18 (72), 79-96. (in Persian).

Allen, E. Larson, c. & Sloan, R. (2009). Accrual reversals, earnings and stock returns. *Journal of Accounting and Economics*, 56 .129-113 ,(1)

Barton, J. & Simko, p. (2002). The balance sheet as an earningsmanagement constraint. *The Accounting Review*, 77 (Supplement).27-1 ,

Bell, T. B., & Carcello, J. V. (2000). A decision aid for assessing the likelihood of fraudulent financial reporting auditing. *A Journal of Practice and Theory*, 19 (1), 169-184. Beneish, M. (1999). Incentives and penalties related to earnings overstatements that violate GAAP. *The Accounting Review*, 74 .457–425 ,(4)

Chen, Y., Chun-Han, Wu., Yuh-Min, C., Hsin-Ying, Li., & Huei Kuen, C. (2017). Enhancement of fraud detection for narratives in annual reports. *International Journal of Accounting Information Systems*, 26 932-45.

Chen, C., & Sennetti, J. (2005). Fraudulent financial reporting characteristics of the computer industry under a strategic-systems lens. *Journal of Forensic Accounting*, 6(1), 23–54.

Chen, S. (2016). Detection of the fraudulent financial statements using the hybrid data mining approach. *SpringerPlus*, 5, 5-89.

Chiu, CC. Lee, TS., Chou, YC., & Lu, CJ. (2002). Application of integrated identification analysis and ANN in data mining. *J Chin Inst Ind Eng*, 19(2), 9–22

Dechow, P. & Dichev, I. (2002). The quality of a ccruals and earnings: The role of accrual estimation errors. *The accounting review*, 77 (1), 35-59.

Dechow, P., Larson, C., and SLOAN, R. (2011). Predicting material accounting misstatements. *Contemporary Accounting Research*, 28 .82-17 ,(1)

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management, *The accounting review*, 70, 193–225.

Ghaderzadeh, S. K., Kordestani, G., & Haghghat, H.(2017) . Identifying and ranking Factors Influencing Levels of Corporate Social Responsibility

Disclosure using Data Mining. *Journal Management System*. 6 (2): 7-46. (In persian).

Graham, J., Harvey, C., Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, 40, 3–73.

Healy, P. (1985). The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics*, 7, 85–107.

Hribar, P., Kravet, T. and Wilson, R. (2010). A new measure of accounting quality. *Review of Accounting Studies*, 19 (1), 506 538

.Ines, A. (2017). The effect of discretionary accruals on financial statement fraud: The case of the French companies. *International Research Journal of Finance and Economics*, 161, 48-62. International Federation of Accountants (IFAC). (2009). International auditing and assurance standard Board (IAASB), Evaluation of misstatements identified during the audit, *International Standard on Auditing*, 450, 368-379.

Jayanthi, S. K., & Sasikala, S. (2013). REPTree classifier for indentifying spam in web search engines. *IJSC*, 3 (2), 498–505.

Jones, J. (1991). Earnings management during import relief investigations. *Journal Accounting Research*, 29 (2), 193–228.

Jones, K. L., Krishnan, G. V., & Melendrez, K. D. (2008). Do models of discretionary accruals detect actual cases of fraudulent and restated earnings? An empirical analysis. *Contemporary Accounting Research*, 25 (2), 499-531.

Kim, Y. J., Baik, B. & Cho, S. (2016). Detecting financial misstatements with fraud intention using multi-class cost-sensitive learning. *Expert Systems with Applications*, 62(15), 32-43.

Kordestani, GH. Tatli, R. (2016). The Prediction of Earnings Manipulation: Development of a Model. *Accounting and Auditing Review*, 23(1), 73-96. (in persian).

Kothari, S., Leone, A. and Charles, W. (2005). Performance matched discretionary accrual measures. *Journal of Accounting and Economics*, 39 (1), 163-97.

Lin, CC., Chiu, AA., Huang, SY., & Yen, DC. (2015). Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. *Knowledge-Based Systems*, 89, 459-470.

McNichols, M. F. (2002). Discussion of the quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review*, 77, 61–69.

Perols, J. L., & Lougee, B. A. (2011). The relation between earnings management and financial statement fraud. *Advances in Accounting*, 27 (1), 39 – 53.

Pourhaidari, O., Samieenejad, N. (2015). Disclosure of corporate governance structure and the likelihood of fraudulent financial reporting. *Applied Research in Financial Reporting*, 4(1), 53-80. (in persian).

Quinlan, J. (1987). Simplifying decision trees. *International Journal of Man Machine Studies*, 27 (3), 221–234.

Rahmani, A., Esmaili, Gh. (2011). The Efficiency of Neural Networks, Logistic Regression & Discriminant Analysis in Defaults Prediction. *Quantitative economics*, 7(4): 151-172. (in persian).

Richardson, S. A., Sloan, R. G., M. Soliman, M. T., & Tuna, I. (2005). (Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics*, 39 (3), 437–485

Report to the Nations on Occupational Fraud and Abuse, Association of certified fraud examiners. (2016). [https://s3-us-west-2, amazonaws.com/acfe-public/2016-report-to-the-nations. pdf](https://s3-us-west-2.amazonaws.com/acfe-public/2016-report-to-the-nations.pdf) Rezaee, Z. and Riley, R. (2010), *Financial statement fraud*, Willy,

New Jersey. Sadgalia, I. Saela, N. & Benabbou, F. (2019). Performance of machine learning techniques in the detection of financial frauds. *Second International Conference on Intelligent Computing in Data Sciences. Procedia Computer Science*, 148, 45–54.

Sajadi, H. Kazemi, T. (2016). A Comprehensive Pattern of Fraudulent Financial Reporting in Iran, *Grounded Theory. Empirical Research in Accounting*, 6, (1), 185-204. (In persian).

Shams al-Dini, K., Daneshi, V., Mahallati Raini, MA. (2018). The Analysis of Benford's Law Ability to Identify Fraud Detection. *Journal of Audit Science*, 18(73), 166-187. (In persian).

Witten, IH., and Frank, E. (2005). *Data mining: practical machine learning tools and techniques –2nd ed.* the United States of America, Morgan Kaufmann Series in Data Management Systems.

Xu, F., & Zhu, Z. (2014). A Bayesian approach for predicting material accounting misstatements. *Asia-Pacific Journal of Accounting & Economics*, 21(4), 349-367.

Zojaji, Z. (2008). Detect penetration in computer networks using cost conscious learning. Faculty of Computer Engineering University of Isfahan. Master's thesis. (In persian).