Financial Statements Earnings Manipulation Detection Using a Layer of Machine Learning

B. Dbouk and I. Zaarour

Abstract—**This paper applies a layer of machine learning technique such as the Bayesian Naïve Classifier (BNC) to enhance the decision making process in the framework of Earnings Manipulation Detection (EMD). It evaluates and competes Manual Auditors' Methods versus a mathematical model in EMD such as the Beneish Model. The Data sets consist of fifty-three (53) Financial Statements acquired from largest corporations over four consecutive years. Using the Beneish model, we classified corporations between manipulators and non-manipulators to establish the training set. The manual audit results for each corporation used to establish a test set as the expert set. In testing for EMD under the mathematical model versus the audit methods, and to evaluate results, a new layer of Machine Learning technique introduced such as the BNC. Our results show that mathematical models outperform auditors. They reveal a classification rate of (86.84 per cent) using the Beneish Model and (60.53 per cent) using Manual Auditors' Methods. Our findings indicate that Manual Auditors' Methods are difficult to detect Earnings Manipulation of Financial Statements. The Main contribution of this research is to use the Machine Learning as a new layer in the Framework of EMD. This approach broadens the scope for auditors, and other financial experts to use Machine Learning with mathematical models through their audit. The results of this study will help regulators and practitioners to detect accounting manipulations and to add value for the auditing, accounting, and financial professions.**

Index Terms—**Beneish model, earnings manipulation, machine learning, supervised classification.**

I. Introduction

Recently, the financial fraudulent behavior continued to increase with financial frauds being detected [1]. The audit committee recognizes that traditional audit methods like manual verification of accounts as payroll, inventory, earnings and old sampling techniques, may no longer be adequate in view of the continuous integration between technological advancements and businesses. Auditors had limited analytic capabilities and jurisdiction to achieve maximum audit compliance with minimum audit costs at in-time judgments. Auditors need new tools in today's massive digital economy of Big Data transactions [2]. Thus, new effective methods are becoming inevitable to complement and strengthen the various audit analytical procedures when running an audit assignment. In this concern, the application of new approaches in EMD such as machine learning have made possible with information technology (IT). The development and availability of IT in auditing, such as Modern (Risk-based) Auditing and Computer Assisted Audit Techniques (CAATS) strongly facilitate the application of machine-learning tools such as the BNC to improve the effectiveness and efficiency of audits [3]. At this level, financial statement information constitute a rich material. It appeared to have predictive abilities in a model of the probability of accounting legislative violations such as the Beneish model introduced and developed in 1997 and updated in 1999 by Professor Messod Beneish. This model could flag and highlight possible areas of concern in financial statements, considering its ability to classify manipulators and nonmanipulators [4].

This research employed the Beneish model to establish a dependable indicator of accounting manipulation for fraud detection procedures by showing the power of such model as compared to normal audit procedures heavily based on Earnings. Its importance for EMD increased when researchers employed it to assess the financial statements in many companies in the USA like ENRON [5]. Moreover, the report of the Association of Certified Fraud Examiners (ACFE) required Certified Public Auditors (CPAs) in 2004 to adopt it during audits when implementing Statement on Auditing Standards No. 99 (SAS 99) to have a reasonable assurance that financial statements are free from material errors or manipulations [6].

In addition to mathematical models, data mining has the capability of filling gaps since it extracts useful information from large data sets. It covers areas of machine learning, pattern recognition, artificial intelligence, and other areas [7]. Reference [8] shows that a main indicator of manipulation can appear by the use of data mining tools in financial information over a certain period. In this concern, the BNC, which is a machine-learning tool, characterized by its transparency, quick training, and high classification power, best fitted this study's EMD classification problem. Thus, we used BNC as a tool of a new machine-learning layer to assess results. Hence, this paper uses a blending mix of mathematical modeling and machine learning. We test the classification power of the Beneish Model model and compare it to the results obtained using the manual audit methods. Unlike prior studies in this concern, we add a new machine-learning technique layer such as the BNC as a tool to assess the Beneish Model and the audit procedure through classification of financial data. M-scores and Manual Auditors' Results are set as the parameters of our constructed BNC network. Our results show that the Beneish Model outperforms manual audit methods. Our classifier revealed a

Manuscript received December 29, 2016; revised March 6, 2017.

Bilal Dbouk is with the Doctoral School of Law, Political, Economic and Administrative Sciences, The Lebanese University (LU), Beirut, Lebanon (e-mail: bilal_dbook@hotmail.com).

Iyad Zaarour is with the Faculty of Economics and Business Administration, The Lebanese University (LU), Beirut, Lebanon.

high classification rate of 86.84 per cent using the Beneish Model and only 60.53 per cent using usual Manual Auditors' Methods. This study will set as an effective complement to the process of Auditing. It will broaden the scope for auditors and other experts when they run an analytical procedure to use advanced mathematical models like the Beneish Model and to extend their procedure by using Machine Learning.

Our Data Sets Consist of Fifty-three (53) Financial Statements of the largest Lebanese corporations in the wholesale of liquid fuel industry, over the years 2006, 2007, 2008, and 2009.

II. Literature Review

A. Traditional vs. Modern Audit Methods

Traditional audit methods include manual account verifications, document vouching by sampling, inventory counts, and use of simple statistics and/or ratios etc.[9]. These methods are retroactive, provide unclear audit reports, and are of limited value in the ever-changing modern business environment because they are slow and backward looking [10]. In addition, they cannot completely fulfill audit verification needs as most of those methods, are done manually with limited sample data [9].

Today, modern audit methods have expected to enhance the credibility of financial statements, and to provide value-added services as well, such as reporting on irregularities, identifying business risks and manipulations, and advising management [11]. Such methods include automatic verifications, data analysis of entire population, system counting, and use of data analysis and/or machine learning – data mining techniques etc. [9]. Modern audit methods referred through literature to continuous audit (CA) or real time audit or risk-based audit or predictive audit as well which all can solve traditional manual audit work problems [9]. Auditors today need to implement these methods and benefit from the acceleration and automation of business IT in order to predict the expected future outcome of process performance at every transaction. Constructing a predictive audit model is now available under these methods by using appropriate data analysis techniques. Hence, the integration with relative mathematical models in the framework of EMD could modernize traditional audit methods. This integration, which is now much easier with Machine Learning, will allow auditors to have a more reasonable assurance about earnings manipulation existence inside financial statements.

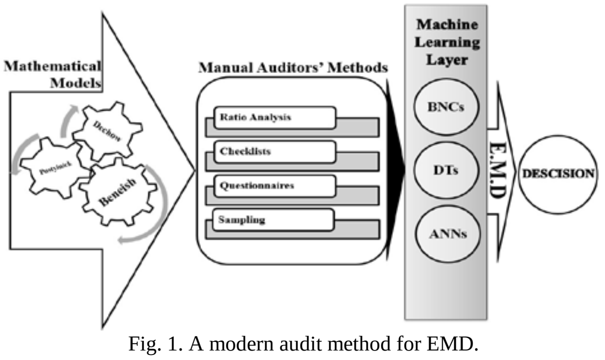

Fig. 1. A modern audit method for EMD.

Fig. 1 depicts a modern audit method in the framework of EMD. This method applies a modern methodology to the audit field. It applies mathematical models supported by machine learning techniques to assist auditors in making a decision about financial statements manipulation detection. Adding a new layer of machine learning techniques such as Bayesian Naïve Classifiers (BNCs), Artificial Neural Networks (ANNs), or Decisions Trees (DTs) after using mathematical models such as of Beneish, Pustylnick, Dechow, etc. at the same level within (represented by the Left Arrow) audit methods will strongly enhance the decision in the framework of EMD.

B. Earnings Manipulation

Accounting manipulations have seen significant growth in academic research during the last decades. According to [12], Earnings Manipulation involves violating the accounting rules and principles. It represents an aggressive earnings management over the boundaries of GAAP and an intervention in the external financial reporting process with the intent of obtaining some private gain. It is an illegal practice and an accounting fraud. GAAP required auditors to focus on Earnings Manipulation activities. However, GAAP sometimes fail due to the flexibility they give to companies in reporting earnings [13].

Literature provides different models to identify manipulation in financial information. The most famous developed models based on accruals accounting and introduced by Healy (1985), De Angelo (1986), and Jones (1991). Reference [14], [15] added a new dimension to the literature of manipulation in financial information. He developed probit and logit model (mathematical model) using a set of different ratios in addition to the accruals which can be used in identifying enterprises applying manipulation in financial information. Other researchers like Spathis (2002) applied logistic regression analysis instead of probit while identifying manipulation in financial information [16]. Reference [17] strongly recommends the use of the Beneish Models by professionals to test for earnings manipulation to prevent earnings managers from deliberately manipulating their financial statements. This research applies the mathematical Beneish Model as an EMD tool and competes it to normal audit procedures using a layer of machine learning technique. This is in an attempt to set up an effective aid to boost and modernize the analytical procedures of auditing.

C. The Beneish Model: An Application Review

Literature include many tools that used in the framework of EMD. Reference [4] review the research that supports further exploration into the use of fraud detection models and tools such as regression analysis, use of nonfinancial information, digital analysis, and ANNs models which all support continuous audit. They show that some researchers applied logistic regression to find that a simple logistic regression model outperforms auditors in fraud risk assessment [18]. Others find that supplementing a checklist with a model helps auditors in assessing fraud risk [19]. Other researchers connected Bankruptcy to manipulations within financial statements to develop a score that is suitable as a sign of financial statements manipulation [20], [5].

In the application of the Beneish model with other models, [21] use forty (40) Italian publicly traded companies divided between Fraudulent Financial Statements (FFS) and Non Fraudulent Financial Statements (Non FFS) over the period 1990-2009 and apply Regression using Jones (1991) model, Lavecker and Richardson's (2004) Modified Jones Model (MJM), and Kothari, Leone, and Wasley's (2005) Model. Unlike the Beneish Model, they find that none of the three models is statistically significant to predict Earnings Manipulation. Similarly, reference [13] uses Binary Logistic Regression to test for Miller Ratio (MR) & Modified Jones Model (MJM) power to detect Earnings manipulation. He finds that both models failed to predict EMD at a statistically acceptable level of confidence on his data with acknowledged manipulation. These findings indicate that the Beneish model has a strong potential in EMD over other models.

Moreover, literature includes prior studies that successfully used the Beneish Model solely and which accordingly verifies the above indication. In this regard, reference [22] uses a sample of seventeen (17) public listed companies who charged by the Securities Commission Malaysia for misstating their financial statements from 1996 to 2014. They find that Beneish model is a reliable tool in predicting potential earnings manipulation in 14 out of the 17 listed companies, which engaged in fraudulent reporting and misstatement. Similarly, Reference [23] apply the Beneish Model and Ratios Analysis for three consecutive years of 2005, 2006, and 2007 as detection tools for Megan Media Holdings Berhad (MMHB). As a result, the authors find that both tools could identify that MMHB involved in manipulating its financial statements. They indicated that auditors might use the Beneish Model to perform audits to have reasonable assurance that financial statements are free from material misstatements.

D. Supervised Classification - Machine Learning Techniques for EMD

References [24] and [25] review extensively the supervised classification techniques of machine learning that involves many applications or algorithms to make computers learn to behave more intelligently by generalizing rather than storing data. The most significant kind of machine learning is Data Mining [24]. Reference [26] defined data mining as "a semi-automatic process that uses statistical techniques, mathematics, artificial intelligence, and machine learning to extract and identify potential knowledge and useful information stored in a large database" [27]. Reference [24] provides that supervised machine learning is the process of learning a set of rules from instances (examples in a training set), or more generally speaking, creating a classifier that can be used to generalize from new instances. In this area of concern, most literature definitions for data mining and supervised classification of machine learning lied within the same concept.

Because people usually hardly find solutions to certain problems in real life, applying machine learning to these problems can improve the efficiency of systems and the designs of machines. According to reference [24], the process of applying supervised machine learning to a real-world problem includes the steps of problem definition, identification of required data, data pre-processing, and definition of a training set, selection of an algorithm, training parameters, and evaluation with a test set. Reference [25] reviews different Supervised Machine Learning techniques for classification and most importantly Statistical Based Machine Learning (Bayesian Belief Networks (BBNs) & Bayesian Naïve Classifiers (BNCs)**.** However, authors concluded that no single learning algorithm might uniformly outperform other algorithms over all datasets. It depends on the type of the classification problem and the accuracy of the applied algorithm that best fits it.

In addition, many researchers have shown the power of BBNs and BNCs in the supervised machine learning techniques applied for a classification problem. A BBN is a graphical model for probability relationships among a set of variables called features [24]. References [28] and [7] have shown that these are powerful tools for knowledge representation and inference under conditions of uncertainty. Reference [29] for instance shows that a BNC represents an effective classification tool that is easy to interpret, and is widely used in banking and financial/claim fraud detection.

In this study, we use the BNC, which is the simplest structure of a BBN [24]. One of the main reasons for using a BNC on a mathematical model such as the Beneish Model is due to being one of the most efficient and effective inductive learning algorithms for machine learning and data mining. Reference [30] study' results indicate that when dependencies between features cancel each other out, there is no influence on the classification when using BNC. That is why it is still the optimal classifier. In this paper, we apply supervised classification using BNC to detect earnings manipulation inside financial statements. This application is justified by the small number of features in an EMD problem supplied by financial ratios of the Beneish Model as attributes for the learning process that is suitable for the application of BNC. Thus, the company's state of manipulation i.e. either a manipulator or nonmanipulator is the classification to distinguish.

III. METHODOLOGY

In this study, we are applying supervised machine learning to a real life problem such as EMD of financial statements. Following [24], the process of applying supervised machine learning to a real-world problem includes the steps of problem definition, identification of required data, data pre-processing, and definition of a training set, selection of an algorithm, training parameters, and evaluation with a test set. In order to reach our objectives for EMD by using a mathematical model and compare it to manual auditors' methods under the application of a supervised machine learning technique, we build our approach through the following steps.

A. Step 1 - Problem Definition

The problem stems from one crucial aspect appeared by the continuous increase of financial fraudulent behavior with more financial frauds being detected. In this part of fraud, Material Misstatements cost economies billions of dollars each year. Another aspect represented by the existence of traditional audit methods like manual verification of accounts and old sampling techniques, which may no longer be adequate in view of the continuous integration between technological advancements and businesses.

B. Step 2 - Period Selection

In this study, we use data extracted from financial statements of four consecutive years starting from 2006 to 2009, with each year ended December 31. Data extracted from the database of a Lebanese Governmental Financial Audit Institution using its audit software. The years of 2006 to 2009 represented the only available closed financial periods of the historical financial statements where the audit institution did the audit.

C. Step 3 - Data Collection

One of the risk factors that may increase the opportunity to commit financial statement fraud is the nature of industry [4]. In this regard, we used the Lebanese liquid fuel industry, since this sector has the largest number of firms that have undergone audit process, based on the financial audit institutions' database. The sample in this study is made of 53 largest corporations in the wholesale liquid fuel industry, with each company representing a separate case. They have approximately the same business size i.e. turnover of more than ten billion Pounds per year (around 7 million US Dollars), and the same legal status of "*Société Anonyme Libanaise*" (S.A.L) or Joint-Stock companies. The data used in this study consist of balance sheets, income statements, and cash flow statements of private firms extracted from one of the largest Lebanese governmental financial audit institutions. It corresponds to the filed firms with the same business size, subjected to an audit by the audit institution in a certain period. The extracted data represent private data extracted from governmental public audit institution. These data were not publicly available due to ethical issues of secrecy of information for not revealing the identity of audited corporations for the public, and according to Lebanese laws, political considerations, privacy, accounts security, and internal governmental practices.

D. Step 4 - Data Preparation

After data cleaning and filtering, 19 only available informative attributes from each of the 53 firms identified and were necessary for applying the Beneish Model. These attributes were Tangible Fixed Assets; Total Receivables; Total Current assets; Total Debt; Total Assets; Net Income; Total Long term Debts; Total Current Liabilities; Working Capital; Total Sales; Total Cost of Goods Sold; Gross Profit; Sales General & Administrative Expenses; Total Depreciation; Total Expenses; Profit or Loss From Operations; NonOperating Income; Results Before Income Tax; and Income Tax Payable.

E. Step 5 - Data Selection

At this level, the above 19 attributes used to derive M-scores for the years 2006, 2007, 2008, and 2009 respectively by applying the formulas of different ratios presented in Table I and the formula of M-score according to the Beneish Model. Thus, eight (8) financial ratios were calculated and constituted the variables of this study. These are DSRI; GMI; AQI; SGI; DEPI; SGAI; LVGI; and TATA. Fifty-three financial statements generated fifty-three (53) times eight (8) dimensional table values for each year data set. The M-score calculations based on t and t-1 time data, for 2006-2007, and 2008-2009 time series respectively. We skipped 2007-2008 series to learn better our training data and to avoid interdependence between the training data set and the test data sets.

F. Step 6 - Data Partition

In this study, we create a partition between two data sets, in which the data extracted from the financial statements of 2006-2007 categorized as training data set, while the data extracted from the financial statements of 2008-2009 used as a test data set. The process of calculating the M-scores for the two data sets help preparing the data for supervised calculations and evaluation using BNC. This process conducted through the application of a computer software tool.

G. Step 7 - Earnings Manipulation Detection

1) The Beneish model (M-score data)

The M-Score model is a mathematical model that uses eight financial ratios to identify whether a company has manipulated its earnings. According to this model, Beneish concluded that a score (M-Score) greater than -2.22 indicates a strong likelihood of a company being a manipulator [14], [15]. The functions and calculation methods of the eight independent variables (financial ratios) that have been determined for this study represented in Table I. They derived from financial statements and created a score to detect earnings manipulation by using the formula of M-score as follows: M – Score = -4.84 + 0.920 DSRI+ 0.528 GMI + 0.404 AQI + 0.892SGI + 0.115DEPI - 0.172SGAI +4.679TATA - 0.327LVGI. Hence, to identify EMD, M-scores calculated for each case by deriving the financial ratios of the Beneish Model to classify companies between a manipulator and nonmanipulator.

In 2006-2007, the Beneish model classified 18 cases as manipulators, 18 nonmanipulated cases, and 17 missing cases due to incomplete data. In 2008-2009, the model classified 6 corporations to be manipulated their earnings, 32 nonmanipulators, and 15 missing cases.

TABLE I: THE BENEISH MODEL RATIOS

No.	Ratio	Name	Formula	Rationale [31]
1	DSRI	Days' Sales in Receivables Index	(Accounts Receivables t / Sales t) / (Accounts Receivables t-1 / Sales t-1)	Shows distortions in receivables that can result from revenue inflation
2	GMI	Gross Margin Index	((Sales t-1 - Cost of Sales t-1) / Sales t-1) / ((Sales t - Cost of Sales t) / Sales t)	Deteriorates margins that predispose firms to manipulate earnings
3	AQI	Asset Quality Index	(1 - (Current Assets t + PPE t) / Total Assets t) / (1 - (Current Assets t-1 + PPE t-1) / Total	Captures distortions in other assets that can result from excessive expenditure capitalization

			Assets t-1)	
4	SGI	Sales Growth Index	(Sales t) / (Sales t-1)	Manages the perception of continuing growth and capital needs predispose growth firms to manipulate sales and earnings
5	DEPI	Depreciation Index	(DE t-1 / (DE t-1 + PPE t-1)) / (DE t / (DE t + PPE t))	Captures declining depreciation rates as a form of earnings manipulation
6	SGAI	Sales, General & Administrative expenses Index	SGA t / Sales t) / SGA t-1 / Sales t-1)	captures decreasing administrative and marketing efficiency through larger fixed SGA expenses that predisposes firms to manipulate earnings
7	LVGI	Leverage Index	((LTD t + Current Liabilities t) / Total Assets t) / ((LTD t-1 + Current Liabilities t-1) / Total Assets t-1)	Shapes earnings manipulation when increasing leverage tightens debt constraints and re-disposes firms to manipulate earnings
8	TATA	Total accruals to total assets index	((WC t - WC t-1) - (Cash t - Cash t-1) + (ITP t - ITP t-1) + (Current Portion of LTD t - Current Portion of LTD t-1) - DE t) / (Total Assets t)	Captures manipulations where accounting profits are not supported by cash profits

Where PPE = Plant, Property and Equipment; DE = Depreciation and Amortization Expense; SGA = Sales, General and Administrative Expenses; LTD = Long Term Debt; WC = Working Capital; ITP = Income Tax Payable; t = current year; t-1 = previous year

2) Manual auditors' methods

For this method, we use the audit results acquired from the audit institution of all cases in the sample during the period 2006 to 2009. They summarize all the manual methods used by an auditor to measure the manipulation in a given case, and mainly earnings in order to classify a case that is likely a manipulator of earnings. Such methods are perceived as traditional because they allow auditors to reach a reasonable conclusion based on a limited sample of a case' parameters. They rely on manual sampling and paper audit analysis such as basic ratio analysis, year-to-year financial data comparison, rather than on advanced mathematical models and computer simulations. We assumed that if the audit institution restates the accounts and mainly earnings figures of a firm, then this firm likely considered as manipulator. This assumption based on the availability and type of data, the absence of clear criteria by the audit institution to classify manipulated companies and to publish their bad status and financial woes for the public. At this level the manual auditors' method revealed 18 manipulated corporations who restated their accounts, 20 non-manipulators, and 15 cases as missing after they were excluded from this test set to match the exact cases in the test set of the Beneish Model (the same 38 cases were tested for more precision in classification).

H. Step 8 - Machine Learning Technique Selection

In this step, to apply supervised machine learning to EMD, as prescribed above, in the process of collecting the dataset, preparing and pre-processing data, selecting features or variables, there exists choosing the right machine learning technique or algorithm, which have to be carried out. In this study, the BNC selected was due to its incremental power in the classification of problems in a framework like EMD. In this concern, our method captures a modern approach as previously represented and prescribed in Fig. 1 in the literature review section. At the level of machine learning layer, a BNC Network type constructed in this study to classify manipulator companies from nonmanipulators. (We assumed the known structure of BNC). The complete dataset of 2006-2007 used as training data for constructing this network. At this level, no findings introduced for the Network, as it is ready for supervised classification through learning and testing.

I. Step 9 - Learning

Learning is a very useful property of BNC. It occurs when an expert builds the network that is refined by learning from data. Vast amounts of data make learning from a dataset more possible and efficient; however, that is usually not the case [32]. In this study, the 2006-2007 Training Data set used for learning the BNC network of the 53 cases by using the Expectation Maximization (EM) algorithm as shown in Fig. 2. Learning of the BNC network of the 53 cases ran using a computer software tool under the EM algorithm. From the advantages of BNC network is that missing values might exist in learning the network due to the causality and the inference nature of the Bayesian networks. The EM algorithm makes it possible to overcome this problem of missing values for the BNC network. This algorithm enables parameter estimation in probabilistic models with incomplete data [33].

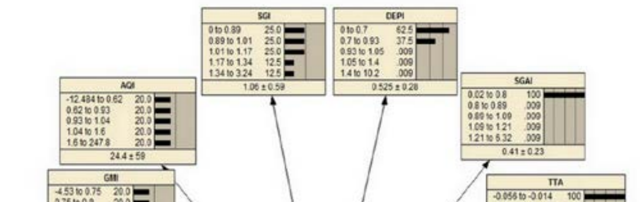

Fig. 2. Learned-training data set using EM algorithm – The BNC Network.

The BNC network shown in Fig. 2 reflects the states of some part of a world (i.e. manipulation of financial statements) modeled and it describes how those states are related. The arrows of the network (also called links) between any two nodes indicate that there are relationships, known to exist between the states of those two nodes. However, the direction of the link arrows corresponds to "causality". The attributes used in the learning process were the prescribed eight financial ratios of the Beneish Model (DSRI; GMI; AQI; SGI; DEPI; SGAI; LVGI; and TATA).

J. Step 10 – Testing, Analysis, and Evaluation

3) Beneish model

After learning the BNC network from the 2006-2007 training data set, we used 2008-2009 test data set based on M-scores to test the learnt classifier in order to evaluate classification results of BNC under the application of the Beneish Model.

4) Manual auditors' methods

At this level, we aimed again to test the detection accuracy of the BNC learnt from the training data set based on M-scores on the test data set. However, the latter now based on manual auditors' results for the audited years covering 2008 and 2009 financial statements, rather than on M-scores.

IV. Findings and Discussion

In any test, there exists the need to measure and evaluate its performance. In this study, we test the performance of BNC over the Beneish model versus over manual auditors' methods. The performance characteristics of a test characterized by its sensitivity, specificity, and threshold. Thus, decision-making would strongly affected by the interpretation of the test results [34].

Machine Learning core's goal is to make a computer generalize information from the done observations. Mostly, performance measurement of a machine-learning algorithm based on its accuracy of classifying a data set [34]. To represent the differing cost of each type of misclassification, a cost matrix can be used (sometimes called a performance or confusion matrix), where each row in this matrix is used to represent the predicted label and each column corresponds to the actual labeling [35]. Machine learning classifiers built to maximize classification accuracy. One method is to use a classifier such as the BNC that provides class probability estimates and compute the expected cost of each label.

In this study, the BNC employed to test with sensitivity, since BNC is an easy classifier to add cost sensitivity to it. BNC will take in the unmodified data set and count all attributes from every instance as normal. It depends on the Bayes theorem and assumes class conditional independence where the effect of an attribute value on a given class is independent of the values of the other attributes to simplify the computations involved and that is why so called a "naive".

Hence, the complete possible states resulted from testing for manipulation existence (M) in all cases (i.e. either a manipulated company or nonmanipulated company) presented by the confusion matrices in Table III and IV respectively. (Predicted=Manipulator denoted by label M, Actual=NonManipulator denoted by label NM).

The training set consisted of 53 cases (corporations) with 18 cases classified as manipulators, 18 nonmanipulated cases, and 17 missing cases, however missing cases was overcome by BNC through the learning process under the EM algorithm using a computer software tool. Two test data sets used for testing. A test data set based on M-score classifications, and a test data set based on manual auditors methods' results. In these sets, M-score predicted 6 manipulated cases, 32 nonmanipulated ones, and 15 missing cases due to incomplete data. However, Manual Auditors Methods' Results predicted 18 Corporations classified as manipulators, 20 cases to be nonmanipulators, and 15 cases (which were missing in the test set of the Beneish Model) excluded to match the same cases of this testing set for more precision. Thus, we tested the same corporations over the Beneish Model and Manual Auditors Methods.

In this regard, each confusion matrix covers the total number of 38 cases in the 2008-2009 test data sets, excluding the missing cases (class labels). These results as prescribed based on the application of a computer software tool.

TABLE II: Confusion Matrices-Naive Bayes Classifier: Beneish M-score

	Actual	Predicted	
	Label	Manipulator	Non Manipulator
M-Score	M	2	4
	NM	1	31

Error rate 13.16%

As shown in Table II above, our approach correctly classified two (2) out of six (6) manipulated companies and thirty-one (31) out of thirty-two (32) nonmanipulated ones with an error rate of (13.16 per cent). This means that in (13.16 per cent) of the cases for which a company committed an Earnings Manipulation, the network predicted the supervised wrong state. This also provides that BNC had a classification rate of (86.84 per cent) for the true positive state of a company to being manipulated its earnings, based on the financial data of M-score. However, based on the assumption that restatement is only due to manipulation, this result largely compared to the manual auditors' methods (actual manipulation) in EMD with (39.47 per cent) error rate (misclassification rate) as shown in Table III below.

TABLE III: Confusion Matrices-Naive Bayes Classifier: Manual Auditors' Methods

	Actual	Predicted	
	Label	Manipulator	Non Manipulator
Manual Auditors' Methods	M	3	15
	NM	0	20

Error rate 39.47%

In this regard, the error classification rate of Earnings Manipulation when using the M-score is three times less than the error produced when using manual auditors' methods. This new finding is promising for auditors and other professionals in evaluating their manual methods or any other scientific model. When an auditor runs an audit assignment, by applying BNC under mathematical models will assist him/her in making a decision about financial statements through EMD.

The Performance matrix for the test of manipulation state (M) which is our interest of quality test, is used as a measure of performance evaluation for our approach in the two testing scenarios (M-Score and Manual Auditors Methods). Sensitivity (type I error) and specificity (type II error) of BNC are used to evaluate classification under M-score versus

Audit Methods as shown in Table IV below.

TABLE IV: PERFORMANCE MATRIX: BENEISH M-SCORE VS. MANUAL AUDITORS' METHODS

	Sensitivity (%)	Specificity (%)
M-Score	33.33	96.88
Manual Auditors' Methods	16.67	100

Sensitivity is the probability that our test will positively give a case with the condition to be a manipulator. It indicates how the test will be true positive (M) in the setting of EMD i.e. (33.33 per cent) for M-score and (16.67 per cent) for manual auditors' methods. While Specificity indicates how a test will be true negative (NM) in companies without having manipulated their earnings i.e. (96.88 per cent) for M-score and (100 per cent) for the specified audit methods. Since, we are interested in EMD; we have more indication towards sensitivity. After constructing the BNC using the financial ratios of the Beneish Model using the computer software tool, and after learning the network, BNC classifies two manipulated cases as manipulators from six such cases correctly (33.33 per cent) using M-score financial data, and therefore produces the best sensitivity.

Our findings in terms of classification accuracy (86.84 per cent) proceed the findings of [16] and [27]. Moreover, our results support the capability of the Beneish Model for EMD. Therefore, introducing a machine-learning layer using BNC under the Beneish Model resulted at most in higher classification accuracy than Manual Auditors' Methods. This supports our approach methodology represented by Fig. 1, to show how machine-learning techniques could assist auditors before making a decision about financial statements in EMD. Adding a new layer of machine learning techniques after using mathematical models such as the Beneish Model competed with and to complement manual audit methods will strongly enhance the decision in the framework of EMD.

V. CONCLUSION

Financial experts mistakenly run analyses when trying to establish relationships between different features. This makes it difficult for them to find solutions to certain financial problems. EMD problem had many concerns in previous research with variation in solutions and limitations. In addition, the identification of material misstatements of financial statements has a critical step in the audit field today, because fraud identification and prevention are more expanding due to the complexity of transactions and growth of global economies in the new era of technology. However, Supervised Machine learning techniques have shown their ability in improving the efficiency of systems and the designs of machines especially in classification [36].

In this study, we used supervised classification as a tool to compete an advanced mathematical model "the Beneish Model" versus "Manual Auditors' Methods" applied to audited financial data over a period from 2006 to 2009. Unlike prior studies, we a machine learning layer using BNC to evaluate both the mathematical model and the auditors' methods. Previous research has shown that BNC is an optimal classifier and it best fits an EMD classification problem. Our Results revealed that the Beneish Model produced the best sensitivity as compared to Manual Auditors' Methods in assessing financial data for detecting earnings manipulation. BNC correctly classified (86.84 per cent) of manipulated companies under the M-score assessment, while it comparably classified (60.53 per cent) under Manual Auditors' Methods of financial data assessment.

Therefore, we can state as a conclusion that traditional manual audit methods, based on simple methods of basic ratio analysis, sampling, manual verification of accounts, need must integrate with technological advancements to comply with the new economy. Modern audit methods through supervised classification and machine learning can assist EMD in the analytical procedures of auditing. For instance, incorporating advanced mathematical models such as the Beneish Model provides better results due to the limited analytical capabilities of auditors in analyzing big financial data and their limited testing for all business cases. To detect earnings manipulation, Machine Learning can strengthen mathematical models. At this level, mathematical models and machine learning push a step forward for machines over human capabilities.

In addition, this study will be contributively valuable towards the potential users of financial statements such as auditors, financial analysts, accountants, government tax controllers, financial forensic investigators and academic researchers as well. It constitutes a preliminary study on machine-learning approach for auditing. It broadens the scope of using mathematical models to identify earnings manipulators and assisting the decision making process, and to predict manipulations under the application of machine learning tools. Furthermore, the comparison in the power of classification in EMD using a machine learning technique like BNC between the Mathematical Beneish Model and the human behavior will encourage auditors and other related investigative parties to search for further tools and techniques to enrich their continuous auditing procedure. They will have better understanding on the attributes of earnings manipulating and machine learning in detecting fraud as well. As a result, they will have a reasonable perception during further studies to replace or integrate limited analytical procedures with such advancements.

ACKNOWLEDGMENT

We express our sincere thanks to the Lebanese Ministry of Finance for their data assistance. We also thank Norsys Software Corp. for using Netica "4.16" free version.

REFERENCES

[1] X. Yue, Y. Wu, Y. Wang, C. Li, and C. Chu, "A review of data mining-based financial fraud detection research," *International conference on wireless communications: Networking and Mobile Computing*, pp. 5519-5522, 2007.

[2] A. Sharma and K. P. Panigrahi, "A review of financial accounting fraud detection based on data mining techniques," *International Journal of Computer Applications*, vol. 39, no. 1, pp. 37-47, 2012.

[3] M. Hematfar and M. Hemmati, “A comparison of risk-based and traditional auditing and their effect on the quality of audit reports,” I*nternational Research Journal of Applied and Basic Sciences*, vol. 4, no. 8, pp. 2088-2091, 2013.

[4] C. E. Hogan, Z. Rezaee, R. Riley, and U. Velury, “Financial statement fraud: insights from the academic literature,” *Auditing: A Journal of Practice & Theory*, vol. 27, no. 2, pp. 231-252, 2008.

[5] M. Mahama, “Detecting corporate fraud and financial distress using the altman and beneish models: The case of enron cor,” *International Journal of Economics, Commerce and Management*, vol. 3, no. 1, pp. 1-18, 2015.

[6] U. J. Nwoye, E. I. Okoye, and A. O. Oraka, “Beneish model as effective complement to the application of SAS No . 99 in the conduct of audit in Nigeria,” *Management and Administrative Sciences Review*, vol. 2, no. 6, pp. 640-655, 2013.

[7] A. Saad, I. Zaarour, and A. Zeinedine, “A preliminary study of the causality of freezing of gait for parkinson’s disease patients: Bayesian belief network approach,” *International Journal of Computer Science Issues*, vol. 10, no. 3, pp. 88-95, 2013.

[8] E. Kirkos, C. Spathis, and Y. Manolopoulos, “Data mining techniques for the detection of fraudulent financial statements,” *Expert Systems with Applications Science Direct*, vol. 32, no. 4, pp. 995-1003, 2007.

[9] S. Kuenkaikaew and A. M. Vasarhelyi, “The predictive audit framework,” *The International Journal of Digital Accounting Research*, vol. 13, pp. 37-71, 2013.

[10] J. Verner. (September 2016). Why you need internal audit at the table. *Published on Business Finance*. [Online]. Available: http://businessfinancemag.com

[11] T. H. Lee and A. M. Ali, “The evolution of auditing: An analysis of the historical development,” *Journal of Modern Accounting & Auditing*, vol. 4, no. 12, pp. 1-8, 2008.

[12] F. Paolone and C. Magazzino, “Earnings manipulation among the main industrial sectors: Evidence from Italy,” *Economia Aziendale Online*, vol. 5, no. 4, pp. 253-261, 2015.

[13] J. Miller, “The development of the miller ratio (MR): A tool to detect for the possibility of earnings management (EM),” *Journal of Business & Economics Research*, vol. 7, no. 1, pp. 79-90, 2009.

[14] M. Beneish, “Detecting GAAP violation: Implications for assessing earnings management among firms with extreme financial performance,” *Journal of Accounting and Public Policy*, vol. 16, no. 3, pp. 271-309, 1999.

[15] M. Beneish, “The detection of earnings manipulation,” *Financial Analysts Journal*, vol. 55, no. 5, pp. 24-36, 1999.

[16] E. Kara, M. Ugurlu, and M. Korpi, “Using beneish model in identifying accounting manipulation: An empirical study in BIST manufacturing industry sector,” *Journal of Accounting, Finance and Auditing Studies*, vol. 1, no. 1, pp. 21-39, 2015.

[17] M. R. Sankar and K. R. Subramanyam, “Reporting discretion and private information communication through earnings,” *Journal of Accounting Research*, vol. 39, no. 2, pp. 365-386, 2001.

[18] T. Bell and J. Carcello, “A decision aid for assessing the likelihood of fraudulent financial reporting,” *Auditing: A Journal of Practice & Theory*, vol. 19, no. 1, pp. 169-184, 2000.

[19] M. Eining, D. R. Jones, and J. K. Loebbecke, “Reliance on decision aids: An examination of auditors’ assessment of management fraud,” *Auditing: A Journal of Practice & Theory*, vol. 16, no. 2, pp. 1-19, 1997.

[20] I. Pustylnick, “Empirical algorithm of detection of manipulation with financial statements,” *Journal of Accounting, Finance and Economics*, vol. 1, no. 2, pp. 54-67, 2011.

[21] E. D'Amico and E. Mafrolla, “The importance of earnings management detection models to identify fraud: A case from italian listed firms,” *Journal of Modern Accounting and Auditing*, vol. 9, no. 1, pp. 68-75, 2013.

[22] M. Kamal, M. Salleh, and A. Ahmad, “Detecting earnings management manipulation by malysian public listed companies: The reliability of Beneish M-Score Model,” in *Proc. Malaysia indonesia International Conference on Economics, Management and Accounting, MIICEMA 2014*, pp. 67-81.

[23] N. Omar, R. K. Koya, Z. M. Sanusi, and N. A.Shafie, “Financial statement fraud: A Case examination using beneish model and ratio analysis,” *International Journal of Trade, Economics and Finance*, vol. 5, no. 2, pp. 184-186, 2014.

[24] S. B. Kotsiantis, “Supervised machine learning : A review of classification techniques,” *Informatica*, vol. 31, no. 3, pp. 249-268, 2007.

[25] S. Anto and S. Chandramathi, “Supervised machine learning approaches for medical data set classification - A review,” *International Journal of Computer Science & Technology*, vol. 2, no. 4, pp. 234-240, 2011.

[26] E. Turban, J. E. Aronson, and T. P. Liang, *Decision Support Systems and Intelligent Systems*, 7th ed. New Jersey: Prentice Hall, 2005.

[27] Tarjo and N. Herawati, “Application of beneish M-score models and data mining to detect financial fraud,” *Procedia - Social and Behavioral Sciences*, vol. 211, pp. 924-930, 2015.

[28] A. Saad, I. Zaarour, P. Bejjani, and M. Ayache, “Handwriting and speech prototypes of parkinson patients: Belief network approach,” *IJCSI International Journal of Computer Science*, vol. 9, no. 3, pp. 499-505, 2012.

[29] C. Elkan, “Magical thinking in data mining,” in *Proc. the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining -KDD ’01*, pp. 426-431, 2001.

[30] H. Zhang, “The Association for the Advancement of Artificial Intelligence and The MIT Press (AAAI) in ‘The optimality of naive Bayes,’” in *Proc. the Seventeenth Florida Artificial Intelligence Research Society Conference*, pp. 562-567, 2004.

[31] M. Beneish, C. Lee, and D. C. Nichols, “Earning manipulation and expected returns,” *Financial Analysts Journal*, vol. 69, no. 2, pp. 57-82, 2013.

[32] D. Heckerman, “A tutorial on learning bayesian networks: Technical report, MSR-TR-95-06,” *Microsoft Research, Advanced Technology Division*, 1995.

[33] C. B. Do and S. Batzoglou, “What is the expectation maximization algorithm?” *Nature Biotechnology*, vol. 26, no. 8, pp. 897-899, 2008.

[34] J. Bradford, C. Kunz, R. Kohavi, C. Brunk, and C. Brodley, “Pruning decision trees with misclassification costs,” in *Proc. ECML-98*, 1998, pp. 131-136.

[35] C. Elkan, “The foundations of cost-sensitive learning,” in *Proc. International Joint Conference on Artificial Intelligence*, vol. 17, vo. 1, pp. 973-978, Lawrence Erlbaum Associates Ltd., 2001.

[36] B. Dbouk and I. Zaarour, “Towards a machine learning approach for earnings manipulation detection,” *Asian Journal of Business and Accounting*, to be published, 2017

Bilal Dbouk is a PhD student at The Doctoral School of Law, Political, Economic and Administrative Sciences, The Lebanese University (LU), Beirut, Lebanon.

He is a Tax Inspector at the Lebanese Ministry of Finance at the LTO - Large Taxpayers Office.

Iyad Zaarour is an assistant professor at the Faculty of Economics and Business Administration, The Lebanese University (LU).

He has an adjacent association with Doctoral School (EDST) of Faculty of science at Lebanese University, and a close collaboration with The Laboratory of Biomedical Department of Islamic University in Partnership with (GREHA) Laboratory of Le Havre University- France. He is associated with the Laboratory (LITIS- EA 4108) at Rouen University, France.