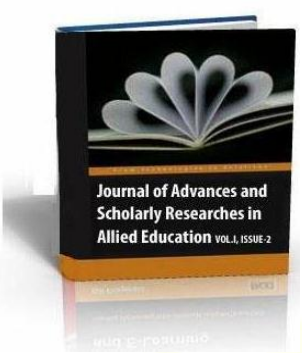
ISSN-2230-7540

A Study on Prevention and Detection of Financial Frauds



Pooja Malik Jain

Asst. Prof., Deptt. of Commerce,

S. A. Jain College,

Ambala City –Haryana (India)

ABSTRACT:

Data mining aims to discover hidden knowledge, unknown patterns, and new rules from large databases that are potentially useful and ultimately understandable for making crucial decisions. It applies data analysis and knowledge discovery techniques under acceptable computational efficiency limitations, and produces a particular enumeration of patterns over the data. The insights obtained via a higher level of understanding of data can help iteratively improve business practice.

Data mining is able to uncover unknown patterns and predict future trends and behaviors in financial markets. It creates opportunities for companies to make proactive and knowledge-driven decisions in order to gain a competitive advantage. Data mining has been applied to a number of financial applications, including development of trading models, investment selection, loan assessment, portfolio optimization, fraud detection, bankruptcy prediction, real-estate assessment, and so on. The competitive advantages achieved by data mining include increased revenue, reduced cost, and much improved marketplace responsiveness and awareness.

ISSN-2230-7540

INTRODUCTION:

Data Mining plays a very important role in prevention of financial statement fraud because the aim of data mining is to find out potential knowledge, unknown patterns and unsuspected relationship from a large set of data. This capability of data mining has been utilized in this research for preventing fraudulent financial reporting.

Data mining tasks can be divided in two subgroups: predictive tasks and descriptive tasks. With predictive tasks, the objective is to predict the value of one attribute, based on the values of other attributes. Due to this nature, predictive data mining along with machine learning is best suited for fraud detection. Predictive tasks make a prediction for every observation. Descriptive tasks however, describe the data set as a whole. It aims to describe the underlying relationships in the data set. This fact accounts for the use of descriptive data mining instead of predictive data mining for fraud prevention. An advantage of the use of descriptive data mining techniques is that it is easier to apply it on unsupervised data. Therefore, this research recommends the use of descriptive data mining techniques for prevention of financial statement fraud.

Descriptive data mining techniques namely association rules, clustering and anomaly detection are suggested as appropriate candidates for prevention of financial statement fraud in this research.

Financial statement fraud is a deliberate misstatement of material facts by the management in the books of accounts of a company with the aim of deceiving investors and creditors. This illegitimate task performed by management has a severe impact on the economy throughout the world because it significantly dampens the confidence of investors. Despite the presence of strong internal control and various internal as well as external audit committees, detecting fraudulent financial reporting fraud is a difficult task when using normal audit procedures due to the following reasons. First, there is a shortage of knowledge concerning the characteristics of financial statement fraud. Secondly, given its infrequency, most auditors lack the experience necessary to detect it.

ISSN-2230-7540

Finally, managers deliberately try to deceive auditors. For such managers, who comprehend the limitations of any audit, standard auditing procedures may prove insufficient. It has also been noted that the increased emphasis on system assessment is at odds with the profession's position regarding fraud detection, since most material frauds originate at the top levels of the organization, where controls and systems are least prevalent and effective. These limitations suggest that there is a need for additional analytical procedures for the effective detection of financial statement fraud.

Cost of financial statement fraud is very high both in terms of finance as well as the goodwill of the organisation and related country. Therefore, this research work discusses various causes and consequences of financial statement fraud.

Detection of financial statement fraud comes in to play only if the prevention mechanism has failed to stop the management in perpetrating the fraud. Therefore, this research work proposes a data mining framework for prevention of financial statement fraud at the first place and detecting it in case of failure of prevention methods.

Data mining methods could possibly assist auditors in detection of fraud, because data mining can use previous instances of fraud to build models in order to identify and detect the risk of fraud. Detection of financial statement fraud is an instance of classification and decision problem and the efficacy of the detection depends on the classification algorithms and the fraud predictors used and how they are combined. Therefore, this research work identifies various financial ratios / variables as input vector to the classification algorithms and further implements three classification algorithms i.e. Decision Tree (CART), Naïve Bayesian Classifier and Genetic Programming for identification and detection of financial statement fraud.

Fraud detection is one of the most important applications of Data Mining. It is widely accepted by both researchers' community and practitioners that there is a requirement of analytical procedures and data mining techniques along with traditional auditing procedures for prevention and detection of financial statement fraud. Auditing firms and

ISSN-2230-7540

procedures are not capable enough to prevent and detect financial statement fraud, since detection of fraud is not their primary objective and auditors have a very little knowledge about the management of the organization. Moreover, standard auditing procedures may prove insufficient because auditors use sampling technique and do not examine each and every transaction. Researchers are working hard for designing and implementing new data mining methodologies for prevention and detection of fraudulent financial reporting. Enormous research work has been already done by many researchers in this area. However, due to upgrading and changing technology and new methods used by management in perpetrating financial statement fraud, there is always a necessity of more research work in this area. This motivated us to work in this area.

REVIEW OF THE LITERATURE:

Prevention and Detection of financial statement fraud should complement and supplement each other. However, Bologna states that prevention should take precedence over detection. Auditing procedures are not skilled for prevention and detection of financial statement fraud, because it is not their key objective. Ethically, management of an organisation is supposed to be accountable for prevention and detection of financial statement fraud but it is almost always accomplished with the consent or knowledge of management.

Failure to detect or prevent financial statement fraud can damage the reputation and the credibility of the audit profession . In order to help auditors, analytical techniques of data mining can estimate the risk of fraud and can help them in understanding the reasons behind the fraudulent financial reporting.

Hence, various techniques of data mining are being used to ease out this extra pressure of prevention and detection of financial statement fraud, from the mind of the auditors.

Data mining is a confluence of the various disciplines such as statistics, artificial intelligence, and pattern recognition. With the coming of data mining as a new field of data

ISSN-2230-7540

analysis, data analyzing techniques can be divided into two groups: reporting techniques and data mining techniques. Reporting techniques refers to the techniques used before, where quantitative and statistical data characteristics are extracted from data and human analysts turn this information into knowledge. These are the techniques currently used in internal control settings. Data mining techniques emphasizes on the semi- automatic process to discover meaningful patterns in large data sets. Especially the data mining characteristic of revealing latent knowledge is very typical and valuable. This characteristic comes forward in the fact that no hypotheses are needed to mine the data, as opposed to pure statistics or data reporting. This is the main reason why these techniques are selected for preventing and detecting financial statement fraud in this research.

Data mining techniques can be used for assisting auditors in prevention and detection of financial statement fraud because these methods are capable of constructing self learned models from historical cases of fraud, which identify and detect the risk of fraud. Data Mining is an iterative process within which progress is defined by discovery of knowledge, which helps in finding the reasons behind financial statement fraud. Data Mining is most useful in an exploratory analysis scenario in which there are no predetermined notions about what will constitute an interesting outcome.

The application of Data Mining techniques for detection and identification of financial statement fraud is a fertile research area. Several law enforcement agencies and special investigative units have used data mining techniques successfully for detection of financial frauds.

Traditional methods of auditing and internal control are not capable enough for prevention and detection of financial statement fraud, because it is a type of management fraud and management is adaptive and usually find easy ways to circumvent the auditing measures. Therefore, several data mining techniques have been implemented by number of researchers for preventing and detecting fraudulent financial reporting. In this research work extensive literature survey is carried out in the area of applicability of data mining

ISSN-2230-7540

methods for prevention and detection of financial statement fraud by focusing on nature of data mining techniques and data specifications.

RESEARCH METHODOLOGY:

After the review of literature, three major areas in usage of data mining methods for prevention and detection of fraud were targeted for future research. These targeted areas include, comparative analysis of four commonly used data mining techniques, identification of financial variables, and designing and implementation of a data mining framework for prevention and detection of fraudulent financial reporting. Seven specific objectives were achieved as a part of this research work.

First objective was achieved by performing thorough study of existing data mining techniques by focusing on purpose, nature of data mining techniques used, data sample specifications and empirical results. Second objective was realised by elaborating the concept of financial statement fraud along with various factors responsible for financial statement fraud. Third objective was accomplished by comparative analysis of four commonly used data mining techniques namely Neural Networks, Decision Tree, Genetic Programming and Bayesian Belief Network on the basis of eight performance criteria. Fourth objective was fulfilled by identifying financial variables from publically available financial statements. These variables were further used as input vector to the proposed framework for realising the primary objective of prevention and detection of financial statement fraud.

Fraud industry incurs a substantial loss to the global economy. Measures to stop fraud from occurring in the first place is termed as fraud prevention whereas identifying fraud as quickly as possible once it has been perpetrated is known as fraud detection. Fraud detection comes into picture once fraud prevention has failed. In practice, of course fraud detection must be used continuously, as one will typically be unaware that fraud prevention has failed.

ISSN-2230-7540

Considering the influence of the loss incurred due to fraud, effective measures and methods should be employed for prevention and detection of financial statement fraud. Data mining methods could possibly assist auditors in reducing the risk of fraud, because data mining techniques are capable of building self learning models by using previous instances of fraud for classifying organisations into fraud or non – fraud. Moreover, data mining methods can design new techniques for preventing and detecting fraudulent financial reporting.

The presence of unexpected deviations or outliers in the financial statement fraud is most likely the indicators of errors, but may also indicate unethical, illegal act of fraud. The perpetrators of financial statement fraud show signs of behavioural characteristics such as living beyond one's means, high personnel turnover. Traditional auditing procedures will not be able to identify these factors and hence unable to prevent financial statement fraud at the first place. In view of the fact that victim organisations are unable to recover their losses, cost effective measures to prevent fraud are critical. The substantial cost of financial statement fraud and considerable threat to the economy of a country inspires us to deduce a data mining framework for prevention and detection of financial statement fraud.

OBJECTIVES OF THE STUDY:

Objective of this research work is to study, introduce, design, apply and evaluate the use of data mining techniques by mainly focusing on two aspects of financial statement fraud – prevention and detection.

Following seven specific objectives are decided and achieved:

Objective – 1: *To study the current use of data mining techniques in prevention and detection of financial statement fraud. This includes the study of the nature of data mining techniques used along with data specifications and empirical results obtained by the use of these techniques.*

ISSN-2230-7540

To achieve this objective existing data mining techniques proposed and implemented by various researchers are comprehensively studied and explored. A review is conducted which includes the analytical and empirical results of various data mining techniques. Studies helped in understanding the applicability, advantages and shortcomings of data mining methods for prevention and detection of financial statement fraud. This also provided path for further research work.

Objective – 2: *To analyse various factors responsible for financial statement fraud. This includes the study of causes and consequences of fraudulent financial reporting.*

To accomplish this objective first concept of financial statement fraud is explored then for better understanding various factors responsible for financial statement fraud are studied followed by the study of effects of fraudulent financial reporting on investors, creditors and a country as a whole.

Objective – 3: *To perform the comparative study of extensively used data mining techniques in identification and detection of financial statement fraud.*

To achieve this objective the four commonly used data mining techniques - Neural Networks, Decision Trees, Genetic Algorithm (GA) and Bayesian Belief Network (BBN) are compared in terms of their performance on the basis of eight parameters on a five point scale ranging from low to very high.

Objective – 4: *To identify financial ratios, variables for prevention and detection of financial statement fraud.*

To accomplish this objective, three financial statements namely balance sheet, income statement and cash flow *statements* are comprehensively studied and analysed. In order to identify financial variables, behavioural characteristics of an organisation along with profitability, liquidity, safety and efficiency are considered. Sixty two financial variables are identified as key input variables to the proposed framework for prevention and detection of financial statement fraud.

ISSN-2230-7540

Objective – 5: *To suggest descriptive data mining techniques for prevention of fraudulent financial reporting.*

To realize this objective three descriptive data mining techniques are suggested to prevent financial statement fraud. In this regard, conventional methods of prevention of financial statement fraud are explored followed by the introduction of three data mining techniques namely, Association Rule, Cluster Analysis and Anomaly Detection.

Objective – 6: *To propose a data mining framework for prevention and detection of financial statement fraud.*

To accomplish this objective, a new data mining framework is proposed in order to prevent fraudulent financial reporting at the first place and detecting it if the management of an organisation is able to perpetrate even in presence of anti – fraud environment. This framework suggests the use of descriptive data mining for prevention and predictive data mining techniques for detection of financial statement fraud.

SIGNIFICANCE:

This research work conducts study and survey of existing use of data mining techniques for prevention and detection of financial statement fraud, which provides better understanding of existing use of data mining technique and their applicability. It helps in setting objectives for conducting the research work. Further whole research work is divided into three parts –comparative analysis of four commonly used data mining techniques, identification of financial variables, and designing and implementation of a data mining framework for prevention and detection of fraudulent financial reporting. Specific findings and significance of each part are given as follows:

First part analyses the data mining techniques commonly used for detection of financial statement fraud. In order to understand the seriousness and cost of financial statement fraud. This part present a brief overview of financial statement fraud, its causes, consequences and various methods of perpetrating it. This part is useful in understanding

ISSN-2230-7540

the concept of fraud especially financial statement fraud along with its causes and consequences. In order to have better understanding the process of fraudulent financial reporting, fraud triangle and a flow chart for financial statement production is also presented in this part. Four data mining techniques are identified as a result of extensive literature survey. Data mining techniques namely Neural Networks, Decision Trees, Genetic Programming and Bayesian Belief Networks were compared on the basis of eight performance criteria. These performance criteria includes: classification accuracy, ease of problem encoding, flexibility, computation complexity, interpretability, optimization capability, scalability and accessibility. Neural Network outperforms the other data mining techniques. This part is useful in selecting a data mining technique for identifying organisations into fraud or non – fraud organisation.

Financial variables for prevention and detection of financial statement fraud have been identified. Sixty two financial ratios and variables were identified on the basis of behavioural characteristics and profitability, liquidity, safety and efficiency of the organisation concerned. These financial variables and ratios are used as input vector to the proposed data mining framework. This study suggests the use of descriptive data mining techniques for prevention of financial statement fraud. Three techniques namely Association Rule, Cluster Analysis and Anomaly Detection were proposed for preventing fraudulent financial reporting. This study is useful in identifying financial variables and selecting a data mining technique for prevention of financial statement fraud.

Third part proposes a data mining framework for prevention and detection of financial statement fraud. This framework suggests the use of descriptive data mining techniques for prevention and predictive data mining for detection of fraudulent financial statements. This framework was implemented by using Association Rule as descriptive data mining method for prevention and three predictive methods namely Decision Trees, Naïve Bayesian Classifier and Genetic Programming for detection of financial statement fraud. This implementation results in seven Association Rules and five Decision Rules for prevention and detection respectively. This part also finds the applicability of the

ISSN-2230-7540

framework by evaluating the performance of the data mining techniques. The performance of classification techniques used in this research has been evaluated by using sensitivity, specificity of the method. This part is useful to both researchers and practitioners in preventing fraudulent financial reporting at the first place and detecting it in case of failure of prevention mechanism.

CONCEPTS AND TERMINOLOGY OF DATA MINING

(a) Exploring Key Terms of Data Mining

Data mining is a process of extraction of useful information and patterns from huge data. It is also called as knowledge discovery process, knowledge mining from data, knowledge extraction or data /pattern analysis. The process of knowledge discovery is explained below:

(b) Knowledge Discovery in Database

The current information age is overwhelmed by data. More and more information is stored in databases and turning these data into knowledge creates a demand for new, powerful tools. Data analysis techniques used before were primarily oriented toward extracting quantitative and statistical data characteristics. These techniques facilitate useful data interpretations and can help to get better insights into the processes behind the data. These interpretations and insights are the sought knowledge. Although the traditional data analysis techniques can indirectly lead us to knowledge, it is still created by human analysts. The current situation however needed a new way to deal with these never ending databases and new methods to analyse this huge amount of data. A new area came into being: Knowledge Discovery in Databases, also known as KDD.

(i) Data cleaning

Real-world data tend to be incomplete, noisy, and inconsistent. Dirty data can cause confusion for the mining procedure. Although most mining routines have some procedures for dealing with incomplete or noisy data, they are not always robust. Instead, they may

ISSN-2230-7540

concentrate on avoiding over fitting the data to the function being modeled. Therefore, a Data cleaning routines attempt to fill in missing values, smooth out noise while identifying outliers, and correct inconsistencies in the data.

(ii). Data integration

The task of data integration combines data from multiple sources into a coherent data store, as in data warehousing. These sources may include multiple databases, data cubes, or flat files.

(iii). Data selection

The task of data selection includes retrieving data relevant to the analysis task from the database.

(iv). Data transformation

Data are transformed or consolidated into forms appropriate for mining by performing the following:

a. **Smoothing**: It works to remove the noise from data. Such techniques include binning, clustering, and regression.

b. **Aggregation**: It means applying summary or aggregation operations to the data. For example, the daily sales data may be aggregated so as to compute monthly and annual total amounts. This step is typically used in constructing a data cube for analysis of the data at multiple granularities.

c. **Generalization** of the data, where low level or primitive (raw) data are replaced by higher level concepts through the use of concept hierarchies. For example, categorical attributes, like street, can be generalized to higher level concepts, like city or county. Similarly, values for numeric attributes, like age, may be mapped to higher level concepts, like young, middle-aged, and senior.

ISSN-2230-7540

d. **Normalization**, where the attribute data are scaled so as to fall within a small specified range, such as -1.0 to 1.0, or 0 to 1.0.

e. **Attribute construction (or feature construction)**, where new attributes are constructed and added from the given set of attributes to help the mining process.

(v). Data mining

Data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining is the process of finding correlations or patterns among dozens of fields in large relational databases. Data mining is a powerful tool because it can provide you with relevant information that you can use to your own advantage.

Data mining is a logical process that is used to search through large amounts of information in order to find important data. The goal of this technique is to find patterns that were previously unknown. It is an essential process where intelligent methods are applied in order to extract data patterns.

SCOPE OF THE STUDY:

This research work focuses on the design and applicability of data mining framework for prevention and detection of financial statement fraud. Data used for implementation of the data mining framework are extracted from publically available financial statements namely balance sheet, income statement and cash flow statements of 114 organizations. Financial variables used in this research are confined to financial ratios available from financial statements; characteristics such as number of outside board members and composition of administrative board are not in the scope of this research work. This research proposes the use of descriptive data mining techniques as opposed to the use of predictive data mining techniques for prevention of financial statement fraud. Text mining algorithms for sentiment analysis of the textual description of financial statements are not in the scope of

ISSN-2230-7540

this research work. The proposed framework considers the use of individual data mining techniques only. Hybrid systems which integrate two or more data mining techniques are not taken in the scope of this research work.

CONCLUSION:

Management should continually evaluate its anti – fraud programs and indicators of fraud for successful fraud prevention. Indicators of fraud are the clues that may warrant further review of a specific area or activity. Some Indicators of fraud are mentioned below :

1. No separation of duties
2. Lack of internal controls
3. High turnover of personnel
4. Unexplained entries or altered records
5. Unusually large amount of payment for cash
6. Inadequate or missing documentation
7. Non – serial number transactions
8. Unauthorised transactions

The above stated occurrences are indicators of fraud which can often lead to opportunities to commit fraud. The incident of fraudulent financial reporting is possible by an employee / manager who have the adequate opportunity and sufficient motive to misstate the numbers in the books of accounts. In order to prevent financial statement fraud, management of an organisation should assess the above mentioned risk factors continuously. Every organisation should implement anti - fraud controls for creating a work environment that values honesty and empower the fraud prevention program. Anti - fraud environment will help the organisations in deterring and preventing fraudulent financial reporting. Anti - fraud control includes the following :

ISSN-2230-7540

1. Establish an appropriate tone at the top.

2. Maintain an adequate and effective system of checks and balances

3. Develop effective corporate governance.

4. Ensure a responsible and accountable board of directors.

5. Hire an objective and independent auditing firm.

6. Establish an independent whistle – blower system.

7. Maintain an independent and effective internal audit.

REFERENCES:

- Nagi Reddy, K. and Ratna kumar, Indian Journal to Applied Psychology Vol. 26 (1).
- Asthana, A.K. Marketing of Banking Services by the Mahasana Urban Co-operative Bank Limited, Mehasana (Gujarat).
- Adrian Payne, The Essence of Service Marketing, Prentice- Hall of India, New Delhi.
- Agarwas, I-I.S. Simple Indian Economics, Lakshmi Narayana Agarwal Educational Publishers, Agra.
- Davar, S.R. Law and Practice of Banking, Progressive Corporation Private Ltd, Mumbai.
- Desai, S.S.M. Rural Banking in India, Himalaya Publishing House, Mumba.
- Dhingra, I.C. Rural Banking in India, S. Chand and Sons, New Delhi.
- Dibakar Panigrahy, Marketing of Plastic Money -A New Avenue of Selling Money, Kanishka Publishers, Distributors, New Delhi-1 10002.
- Frank Jekins, Advertising, 3rd Edition, Macrnillan India Ltd, Delhi.

ISSN-2230-7540

- Gandhi, J.C. Marketing - managerial, Tata Mc Grew Hill Publishing Co Ltd. New Delhi.
- Geoffery Crowther, An outline of money, Universal Book Stall, Delhi.
- Gordon, E. Natarajan, Banking -Theory, Law And Practice, Himalaya Publishing House, Mumbai. 2002.
- Gulshan, S.S. Gulshan, K. Kapoor, Banking Law And Practice, S.Chand & Company LTD, New Delhi.
- Gupta, C.B. An introduction to statistical methods, Vikas Publishing House. Pvt. Ltd. New Delhi.
- Hellen Woodruffe. Service Marketing, Macmillan India Ltd. Delhi.
- Ian Chaston, New Marketing Strategies, Response Book, A division of Sage Publications, New Delhi.
- Jagwant Singh, Indian Banking Industry-Growth and Trends in Productivity, Deep &Deep Publication, New Delhi.
- S.M. Bank Marketing, Himalya Publishing House, Mumbai. .
- S.M. Marketing Management in Indian Perspective, Himalaya Publishing House, Mumbai.
- S.M. Service Marketing, Himalaya Publishing House, Mumbai.
- Jhingan, M.L. Money, Banking, International Trade and Public Finance, Vrinda Publication (p) Ltd.