IMPLEMENTATION OF DATA MINING ALGORITHMS IN DETECTION OF FRAUD IN BANKS: EVIDENCE FROM NIGERIA.

Ugwu Ikechukwu Virginus Ph.D.

Department of Accountancy Chukwuemeka Odumegwu Ojukwu University (COOU), Igbariam Anambra State, Nigeria virginusugwu418@gmail.com

ABSTRACT

This study was embarked upon to determine the implementation of data mining algorithms techniques in fraud detection in Nigeria banks. To achieve this, the opinion of banks' staff: (internal auditors, internal controls and accountants), were sort on how implementation of the three explanatory variables of data mining algorithms of Classification, Clustering and Neural Networks applies on fraud detection in Nigeria banks. The study, selected 15 banks comprising a population record of 668 staff and 400 sample sizes that was obtained using Taro Yamane's formulae. A survey research design was adapted and questionnaire was administered to the sample size to obtain data for this study. The methodology used Likert Scale rating on the three research questions; and its suitable descriptive statistics, while ANOVA test statistics was used to test the hypotheses formulated at 0.05 level of significant. The findings of this study from the agreement of the respondents indicated that implementation of data mining algorithm techniques of classification, clustering and neural networks detects fraud in Nigeria banks. The main recommendations are that management and stakeholders of banks and other financial institutions are encouraged to implement algorithms techniques in fraud detection. They should also train their control staff in the use of data mining techniques.

Key Words: Data Mining Techniques, Algorithms-Classification, Algorithms-Clustering, Algorithms- Neural Networks, Fraud Detection, Banks.

Introduction

Literatures have shown that there are three main monitoring mechanisms that have been identified in the corporate governance participation in banks fraud detection. These are external auditing, internal auditing, fraud auditors and directorships, (Anderson, Francis & Stokes, 1993; Blue Ribbon Committee, 1999; Institute of Internal Auditor, 11A, 2003; Coram, Ferguson & Moroney, 2011). In this paper we consider the opinions of the monitoring mechanism in the implementation of data mining algorithms in banks fraud detection.

Fraud detection or control is viewed as a relatively recent corporate function, (Dionne, 2013). The modern control started after 1955 and up-to 1970, the concept covered more on insurance market and through that it developed to complement several other business activities, (Dionne, 2013). The necessity for fraud detection or control has come up based on the changes on the strategic corporate/business operating environment and the expansion in transaction volumes that have as well affected the methods corporations approach, and the way they handle fraud, (WCO, 2010). Every business increase has a corresponding increase of uncertainties that drives corporate management to desire a better structured and systematic way to handle the imminent fraud. It is through fraud control that corporations address the increasing demand of the modern business operating environment and the consequent threshold fraud by endeavouring to address the fraud whenever they are imminent.

It was the issues of risks of fraud that brought about the necessity of the word "fraud" control and detection measures that covers numerous types of inevitable "Loss" caused by business environment, technology, humans, organizations and politics.

The word Fraud is an activity that takes place in a social setting and has severe consequences for the economy, corporations, and individuals (Silverstone & Sheetz, 2007). History of fraud dated from the cradle of business. The South Sea Bubble of 1720 is the best known early episode of fraud. The company formed in England in 1711 to trade with Spanish America, was allowed in 1720 to assume responsibility for Britain's National debt in return for a guaranteed profit. This complicated arrangement ignited a speculative boom with unscrupulous financiers that took advantage of the public excitement about assured profits to form other companies with dubious intentions. Many of the newly formed companies, some of which sought to extract gold from sea water soon failed together with the south sea company leaving thousands of shareholders to lose their investment. This caused financial catastrophe in London, Paris and Amsterdam. Subsequent investigations revealed fraud and corruption among ministers some of whom resigned and some committed suicide (Onogun, 2009; Adedeji, 2005).

Progressively, the Dictionary of Economics and Commerce confirmed that 200 banks failed in England alone between 1815 and 1850 just within a period of 35 years, one of the reasons attributed to the failure human inadequacy and improper fraud detection, (0wolabi, 2010).

Lack of fraud detection failure is the determining character of the global financial crisis (Witoonchart, 2012). In an effort to reduce fraud which culminated to global financial crisis, (Witoonchart, 2012) stated that thirty five (35) official Anti-Fraud and regulatory bodies have been formed and recognized internationally to regulate, supervise and investigate people and organizations activities (Davis, 2011).

The problems of frauds are the same even in Nigeria and have had sever negative consequences on the country and its global image. Lack of Fraud detections have caused instability in the Nigerian economy resulting to a high mortality rate of business organizations and the consequent losses of revenue, huge financial losses to business organizations and their customers, depletion of shareholders funds and capital base as well as loss of confidence in business investment (Hamilton & Gabriel, 2012).

In Nigeria alone, several legislations were put in place to reduce and to alleviate and if possible to eradicate the occurrence and incidences of fraud in the industry (Awolabi, 2010). Most popular and prominent among them are: (Company and Allied Matters Decree No 19. 1990 (CAMD 1990), now CAMA, Declaration of Asset Act 1990, National drug law enforcement Agency Act 1990; Special Tribunal (miscellaneous offences) Act 1990; The Central Bank of Nigeria (CBN) Decrees No 24 of 1995; The Nigerian deposit Insurance Corporation Decrees No 22 of 1998; The Bank and other financial Institution Decree (BOFID) 1999; Economic and Financial Crime Commission Act 2004; (CBN) Prudential guideline for Deposit of Money in Banks in Nigeria; Money Laundry Act and so on.

Statement of the Problem

But in all the measures that have been put in place for fraud detection both in Nigeria and internationally, frauds have continued. The PwC global economic crime survey of 2018 found that half of the 7,200 companies they surveyed had experienced fraud of some kind or another. Several efforts have been taken to detect fraud globally. Investors have believed that the presence of external auditors could prevent fraud. But, the truth is that, whatever the size of the organization, external audit is terribly bad at fraud detection and the scope of their responsibility do not cover fraud detection. One of the survey by (Pricewaterhouse Coopers, 2011), showed that perhaps only about 2 percent of frauds were detected through external auditor (Taylor, 2011).

Historically, management believed that external auditors would uncover fraud but the emergence of Sarbanes–Oxley specifically holds management responsible for fraud detection (Loftus, 2011). Yet, evidence from the literatures had shown that, fraud had continued undetected using the routine activities of external and internal auditors.

Several works have been done on fraud management, control and detection both in Nigeria and elsewhere. In Nigeria we identified these authors, (Badara & Saiden, 2014; Salamu & Agbeja, 2007) and outside Nigeria, we identified these authors, (Omar & Baker 2012; Endaya & Hanefah; Thenfanis., Drogalis & Giovani, 2011; Domenic & Nonna, 2011; Intakhan & Ussahawanitchakit, 2010; Feizizadeh, 2012; Collier, Dixon & Marston 1991; Farcane, Blidset & Popa, 2009; Mui 2009; Stribu et al, 2009) but none of these works made use of data mining algorithms in fraud detection. Therefore, this current study want to see how the implementation of data mining algorithm could detect fraud in banks.

Objective of the Study

The main objective of this study is to find out on how the implementation of data mining algorithm could detect fraud in Nigeria banks. Other specific objectives are to find out the opinions of the three monitoring mechanism (auditors, internal control and accountants) on how the implementation of:

1) Data mining algorithm-Classification techniques could detect fraud in Nigeria banks;
2) Data mining algorithm-Clustering techniques could detect fraud in Nigeria banks;
3) Data mining algorithm-Neural Networks techniques could detect fraud in Nigeria banks.

Research Questions

1) To what extent does the implementation of algorithm-classification detects fraud in Nigeria banks.
2) To what extent does the implementation of algorithms-clustering techniques detects fraud in Nigeria banks.
3) To what extent does the implementation of algorithms neural networks detects fraud in Nigeria banks.

Hypotheses of the Study

H_{01}: There is no significant agreement among the opinions of fraud auditors, internal auditors and accountants on how the implementation of algorithms classification techniques could detects fraud in banks;

H_{02}: There is no significant agreement among the opinions of fraud auditors, internal auditors and accountants on how the implementation of algorithms clustering techniques could detect fraud in banks;

H_3: There is no significant agreement among the opinions of fraud auditors, internal auditors and accountants on how the implementation of algorithms-Neural Networks techniques could detect fraud in banks.

Review of Related Literature

Data Mining Algorithm implementing on detecting fraud

The concept of implementation of data mining algorithm in preventing and detecting fraud came up recently as we are witnessing a massive increase in the quantity of data (text, pictures, audio, video etc.), both at global and economic level entities. This process is amplified by the entry of any of these mentioned above into the virtual environment. Data comes from everywhere, from numerous and diverse sources like contracts, customer interactions, call centres, social media, phones, emails, faxes, and others. The trend is to use these data for the interest of the entity (conceiving strategies, opportunities identification, goodwill development, preventing and detecting fraud etc., (Adrian, 2015; Bolton and Hand, 2001).

The use of data analysis processes and the software dedicated to these operations provide extensive and in-depth analysis of the phenomena and processes of the informal economy, fraud and corruption, as the information and communication technology becomes an instrument of registered (formal) economy. Although on the analytical market, there is a wide spectrum of specialized tools capable to support and enhance the antifraud activity. Some established survey reports indicate that corporation managers are not taking advantage of implementing data mining algorithms in preventing and detecting fraud in their corporation.

A survey carried out by (Adrian, 2015) found out that implementation of forensic data analytics tools are currently in use in the organizations, but there is much lower adoption of more sophisticated forensic data analyses tools. The outcome of the result depicted only 1. 65% of survey participants that report the use of spreadsheet tools such as Microsoft Excel and 43% report the use of database tools such as MS Access or MS SQL Server. While these tools are important to every forensic data analyses program, they often focus on the matching, grouping, ordering, joining or filtering of data that is primarily descriptive in nature. Implementation of data mining algorithm in fraud detection is so much needed especially in banks where large transactions are carried out spontaneously from different sources at a time. In the same vein fraudulent transaction may accompany the daily large operation that requires detection with data algorithms.

Fraud Detection

The definition of fraud in Concise Oxford Dictionary is "criminal deception; the use of false representations to gain an unjust advantage." Fraud is as old as humanity itself and can take an unlimited variety of different forms. However, in recent years, the development of new technologies (which have made it easier for us to communicate and helped in increasing our spending power) has also provided yet further ways in which criminals may commit fraud. Traditional forms of fraudulent behaviour such as money laundering have become easier to perpetrate and have been joined by new kinds of fraud such as mobile telecommunications fraud and computer intrusion, (Deshpande. Siddiq, Alam and Parmars 2016)

Fraud detection can be seen as a set of activities undertaken to prevent money or property from being stolen or obtain2ed through false pretences. Fraud detection is applied to many industries such as banking or insurance. In banking, fraud may include forging checks or using stolen credit cards. Other forms of fraud may involve exaggerating losses or causing an accident with the sole intent for the pay-out. There is a great difference between fraud prevention and fraud detection. Fraud prevention describes measures to stop fraud from occurring in the first place (Bolton and Hand, 2001), and fraud detection comes into picture where fraud prevention fails.

With an unlimited and rising number of ways that people commit fraud, detection can be difficult to accomplish. Activities such as reorganization, downsizing, moving to new information systems or encountering a cyber-security breach could weaken an organization's ability to detect fraud. This means techniques such as real-time monitoring for frauds is recommended. Organizations should look for fraud in financial transactions, location, devices used, initiated sessions and authentication systems (Colleen McCue, 2015).

Fraudsters normally work individually or in group. The trend in fraudulent activity can change depending upon the situation as we see the number of frauds rising during recession or during festive seasons etc. Fraud occurs in all aspects of human life, but the motivation may not be always the same. Here we can think of motivation behind fraud as money and/or power. If motivation is money then it can lead to different types of fraud such as banking fraud, telecommunications fraud, insurance fraud, forgery fraud etc.

Data Mining in Fraud Detection

Kamlesh, Diwakah and RamMilan, (2019), defined data mining as the process of discovering interesting knowledge such as associations, patterns, changes, anomalies and significant structure from large amounts of data stored in a databases, data warehouse or other information repository. Data mining techniques in fraud risk management is a process which finds useful patterns from large amount of data in account. It is a logical way that is applied to search through huge data in order to extract useful information for decision purposes, (Bharati, 2010).

Data mining is a set of computer- assisted techniques designed to automatically mine large volumes of integrated data, and also for new hidden or unexpected information, or pattern. It also supports customer relationship and fraud detection, Elena, et al, 2014). Data mining is the process of extracting knowledge hidden in large volumes of data. Data mining tools search for trends or anomalies without knowledge of the meaning of the data but the anomalies may not necessarily be an indication of fraud but can be the result of a range of different factors, (Denker, 2003). Control auditors can use data mining tools and techniques to examine the entire population of transaction in order to select samples for test controls and identify fraud. Corporations need to have access to the data and software tools as well as the techniques and knowledge necessary to make intelligent use of vast amount of financial and non-financial information in the aim of detecting and stopping fraud, (Elena, Andrei and Lucian, 2014).

There are things to note when there are anomalies in data. Denkar, (2003), observed that in many case they are as a result of faulty data entry, whereby the user has typed in one value instead of another. Sometimes errors can be as a result of software or hardware malfunctions, resulting in corrupt data but sometimes, it could be fraud.

Fraud audit is defined as a set of audit procedure performed over a business transaction population in order to increase the likelihood of identifying fraud. Thus using data mining in fraud detection is the process of obtaining and analysing transactional data to identify anomalies or patterns indicative of a specific fraud scheme, (Vona L, 2008).

Objective of use of data mining in fraud detection is to find a discrete number of transactions that can be examined using fraud audit procedures. Data mining algorithms fraud detection final purpose is to identify one fraudulent transaction and afterwards have the plan dictate how the sample containing the transaction will be extended. There are various data mining tools and techniques that can be applied to identify transactions consisted with a specific fraud scheme, (Vona L, 2008).

Data Mining Algorithms in Fraud Detection

We can also define data mining as a process of discovering previously unknown patterns in the data using *automatic iterative methods*. Algorithms are iterative step-by-step procedure to transform inputs to output. The application of sophisticated algorithms for extracting useful patterns from the data differentiates data mining from traditional data analysis techniques. Most of these algorithms were developed in recent decades and have been borrowed from the fields of machine learning and artificial intelligence. However, some of the algorithms are based on the foundations of Bayesian probabilistic theories and regression analysis, originated hundreds of years ago. These iterative algorithms automate the process of searching for an optimal solution for a given data problem, (Eric Conrad and Joshua Feldman, 2016)

Algorithm is a mathematical process to solve a problem using a finite number of steps, an algorithm is the set of instructions that defines not just what needs to be done but how to do it. A bank loan officer may want to analyse the data using data mining algorithm in order to know which customer is risky or which are safe. Data mining and predictive analytics represent effective approaches to addressing this pattern of illegal behaviour. Specifically, modelling algorithms that incorporate clustering techniques and anomaly detection can be used to identify patterns of behaviour or activity that deviate from established patterns and trends. Data diverging or deviating from "normal" can be identified for further evaluation. On the other hand, rule induction models capitalize on the fact that people frequently are not creative or unique when they commit fraud (Colleen McCue, 2015),. Although there are important individual differences in this type of criminal behaviour, the secondary gain or desired goal generally structures the approach somewhat, which may limit the options for committing fraud. Therefore, rule induction models can be used to characterize and model known patterns of fraudulent behaviour that can be applied to new data in an effort to quickly identify these patterns. Finally, the use of integrated approaches that utilize both scoring algorithms and unsupervised learning models can allow the analyst to exploit knowledge regarding previously identified or otherwise known or suspected patterns of criminal behaviour, while remaining open to discovering unknown or unanticipated patterns of suspicious behaviour. This combined approach of confirmation and discovery represents one of the more powerful aspects of data mining algorithm, (Ali Safa Sadiq and Kayhan Zrar Ghafoor, 2019).

Algorithms builds a mathematical model based on sample data, known as training data in order to make predictions or decisions without being explicitly programmed to perform the task, (Friedman, 1998). It uses, machine learning anomaly detection in data mining, detect anomaly known as outlier detection. It is the identification of rare items, events or observations which raise suspicions by differing significantly from the majority of data, (Zimek, Schubert, 2017). A typical anomaly items represents an issue such as bank fraud, structural defect, medical problems or error in a text. Anomalies are referred to as outliers, novelties, noise, deviations and exceptions, (Hudge and Austine, 2004).

Bank Fraud

Bank fraud can be defined as an unethical and/or criminal act by an individual or organization to illegally attempt to possess or receive money from a bank or financial institution. Let's take a look at several types of bank fraud which exist. There are bank fraud such as accounting fraud which manifest as demand draft fraud, remotely created check fraud, bill discounting fraud, duplicating or skimming card information, check kiting, forged or fraudulent documents, forgery and altered cheques, fraudulent loan applications, fraudulent loans, empty ATM envelope deposits, stolen ATM, fictitious bank inspector or staff, identity theft or impersonation, money laundering, payment card fraud, booster cheques, stolen payment cards phishing internet fraud, prime bank fraud, stolen cheques etc.

Credit card fraud: One of the primary methods used to perform credit card fraud is the act of duplicating or reproducing the information located on the magnetic strip of the card. This illegal process is known as skimming. Criminals can also perform credit card theft by adding a skimmer of their own on top of the original in an effort to illegally utilize the card and its confidential information. One of the more common forms of credit card fraud occurs after having a debit or credit card stolen or lost. In these situations, an unauthorized party has access to another individual's credit or debit card numbers (although not knowing the PIN number will make it virtually impossible to withdraw monetary funds from an ATM). In other scenarios, credit card fraud may be performed by retailers and merchants who duplicate the information while they have the card in their possession during a purchase.

Illegal Check: The term check fraud refers to illegally using a check for unauthorized financial gain. There are several ways that check fraud can occur and here are a few examples: depositing a check into account without the proper authorization; altering a check by changing bank information such as account numbers; using a check to make a payment knowing that there are insufficient funds in the account; altering the payment amount on a check; using checks for false invoices.

Algorithms-Classification Techniques in Fraud detection

Algorithms classification is one of the Data Mining Techniques that is used to analyse a given data set and takes each instance of it. This assigns the instance to a particular class in such that classification error will be least. It is used to extract models. Algorithm classification, define important data classes within the given data set. According to (Aggarwal, 2015), classification is a two-step process. During the first step, the model is created by applying a classification algorithm. That is on training data set. Then in the second step, the extracted model is tested against a predefined test data set. That is to measure the model trained performance and accuracy. So classification is the process to assign class label from a data set whose class label is unknown. Classification is, therefore, referred to as supervised learning because an example data set is used to learn the structure of the groups, just as a teacher supervises his or her students towards a specific goal. While the groups learned by a classification model may often be related to the similarity structure of the feature variables, as in clustering, this need not necessarily be the case. In classification, the example training data is paramount in providing the guidance of how groups are defined. Given a data set of test examples, the groups created by a classification model on the test examples will try to mirror the number and structure of the groups available in the example data set of training instances, (Aggarwal, 2015). There are algorithm classification analyses:

Classification Analysis: It is a supervised based data mine algorithm. Classification is a systematic process of grouping the similar data into different classes or identifying to which of a set of categories different types of data on the basis of structures using discrete function and obtaining relevant information about data and metadata.

Decision tree induction classification algorithms: In decision tree induction algorithms is suitable for analyse and categories the big data. Decision tree classifiers are useful for break a more complex decision into a collection of the simpler decision. In decision tree structure all internal node represents a test on an attribute, all branch represents a result of the test, each leaf node represents a class label and topmost node in a tree is the root node.

Evolutionary-based classification algorithms: Evolutionary algorithms used for selecting proper data for analysis in good optimization solutions and solution of the multi-objective problem. There are different types of evolutionary algorithms such as genetic algorithms, evolution strategies,

evolutionary programming and so on. Genetic algorithms were mostly used for mining classification rules in large datasets, Patil et al 2006) proposed a hybrid technique using for genetic algorithm and decision tree that improving the efficiency and performance of computation

Algorithm Clustering Techniques in Fraud Detection

It is a technique of partitioning a set of data into clusters or groups of objects. The clustering is done using algorithms. It is a type of unsupervised learning as the label information is not known. Clustering methods identify data that are similar or different from each other, and analysis of characteristics is done. Cluster analysis can be used as a pre-step for applying various other algorithms such as characterization, attribute subset selection, etc. Cluster Analysis can also be used for Outlier detection such as high purchases in credit card transactions.

Aggarwal, (2015), argued that it is an unsupervised based data mine algorithm. In other words, clustering is a process of grouping similar objects into classes. The cluster is a collection of data objects that are similar to one another within the same cluster and are dissimilar to the objects in other clusters. Clustering analysis is the process of identifying data sets that are similar to each other to understand the differences as well as the similarities within the data, (Jiang and Yang, 2016). It is one of the techniques that could apply in banks fraud detection as many transactions are effected on the daily bases. These can be categorized into partitioning methods, hierarchical methods, density-based methods, grid-based methods, model-based methods and constraint-based methods (Jiawei, Micheline and Jian, 2011).

Partitioning based clustering algorithms: In this approach, large data sets are divided into a number of partitions known as *k* partitions, where each partition represents a cluster they known as K-mean, (Jiang and Yang, 2016). J. C. Bezdek et al proposed Fuzzy- C Mean's approach using K-means technique for distributed large dataset (Jiang and Yang, 2016).

Hierarchical based clustering algorithms: In this approach, large data are organized in a hierarchical manner based on the medium of proximity. First initial node is called root cluster which can derive several child clusters. It follows a top-down or bottoms up a strategy to represent the clusters. T. Zhang et al proposed to Birch algorithm using hierarchical clustering which handles streaming data in real time and extracting semantic content was defined in Hierarchical clustering for concept mining, (Hinneburg and Kein, 1999).

Density-based clustering algorithms: Within this approach, clusters are formed based on the data objects regions of density, connectivity, and boundary. Each cluster grows in any direction based on the density growth (Arun and Jabasheela, 2014), while A. Hinneburg et al proposed Den clue algorithm using density based algorithms which can handles, separating on a different type of data and mining large amount of data (Hinneburg and Kein, 1999).

Grid-based clustering algorithms: As obtained in this algorithm, clusters are divided into a number of grids for fast high processing, while (Hinneburg and Kein, 1999), proposed OptiGrid algorithm which handles terabytes volume data and its according to this approach cluster is defined as a finite number of a cell that forms a grid structure in function

Algorithms Neural Network Techniques in Fraud Detection

Algorithms neural network is a set of connected input or output units and each connection has weight present with it. Algorithms Neural Networks works as a set of connected input and output units and

each connection has weight present within it. When it is connected, at this time, during the learning phase, network learns by adjusting weight in order to be able to predict the correct class label of the input tupelos. There is this remarkable ability that algorithm neural networks has to drive meaning from a complicated or imprecise data and can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques, (Ramageri, 2014). A neural network is a series of algorithms that endeavours to recognize underline relationships in a set of data through a process that mimics the way human brains operates. Chyan-long, (2018), confirmed that neural network is a structure similar to the neurons in a human brain. It is an information processing system that mimics biological nerves and that can receive and combine multiple inputs to make predictions. Neural networks can adapt to changing input; so the operations generates the best possible result without needing to be redesign for the output criteria. There is artificial neural networking that serves as a type of artificial intelligence machine where the mathematical method is used to make the computer to have ability to deduce the outcome through the computer's rapid calculation ability, (Chyan-long, 2018). There are several types of neural network. But the common one is feed-forward perception. This network consists of three layer nodes. These are the input layer, the hidden layer and the output layer and within this, data passes forward through the network. During this process, a transaction presented to the input will result in a score at the output, which can be used to tag the corresponding transaction as suspicious fraudulent or legitimate. When a neural network produces a score it is very difficult to understand why it produces the result it did (Poonam, 2016). Algorithm Neural network can be very good at dealing with highly skewed data like fraud card data. It was stated by Kilos, (2006), that the purposes of any fraud detection system is to produce a score that reflects the probability of fraud given the evidence. According to Rajdeepa and Nandhitha, (2013) algorithm neural networks are well suited for continuous valued input and output fraud detection.

Theoretical Framework

In the theoretical framework of this study we state that there are causal factors that are considered to remove or deter fraud and these are best described in the Fraud Triangle. This was first propounded by (Donald R. Cressey 1951). The Fraud Triangle describes three main factors that are present in every situation of fraud. These factors are: 1) Motive or pressure i.e. the need for committing fraud which might be the need for money, power, wealth etc.; 2) Rationalization i.e. the mind-set of the fraudster that justifies them to commit fraud; and; 3) Opportunity i.e. the situation that enables fraud to occur and this happens often when internal control and fraud control mechanisms are none functional in an organization. The only means left for such an institution is to break the Fraud Triangle. Breaking the Fraud Triangle is the key to fraud deterrence. Breaking the Fraud Triangle implies that an organization must remove one of the elements in the fraud triangle in order to reduce the likelihood of fraudulent activities. Among the three elements of fraud triangle, it is only the removal of "Opportunity" and it is this very element that is most directly affected by the system of management internal controls mechanisms and this generally provides the most actionable route to deterrence of fraud, (Chaudhary, 2013).

Empirical Review

Uğurlu and Sevim (2015) studied the importance of financial statement accuracy to credit risk management in banks. The purpose of the study was to predict fraud risk to prevent bank credit risk. The authors used an artificial neural network (ANN) methodology to analyse 289 organizations for the year 2007. The organizations for the study were selected based on ability to fulfil obligations, which could lead to Type 1 and Type 2 errors. The authors identified common features of fraudulent financial statements, 64 such as small net assets and rapid growth. However, the findings indicated that there is a lack of generally accepted variables for detecting fraud, due to the differing reasons for the fraud.

Although many models exist for detecting fraud, the authors found the ANN model to have 90% accuracy for detecting financial statement fraud.

Albashrawi (2016) provided a review of the various studies on data mining techniques for fraud detection from 2004 to 2015. The author found 41 techniques across 65 published articles. The largest application of data mining techniques regards financial statement fraud and bank fraud. Common detection methods of the most used techniques are outlier identification and hidden trends. The work of Albashrawi (2016) helped to identify the best methods of data mining detection based on the type of fraud, frequency of use, and accuracy. The findings indicate that the logistic regression model is used most frequently, and that supervised techniques outperform unsupervised techniques for detecting financial statement fraud. For the purpose of this research, financial fraud is classified as financial statement fraud, bank fraud, insurance fraud, and other fraud, with financial statement fraud and bank fraud making up 63% of the total found in the various 61articles on data mining. Albashrawi (2016) noted that one third of the articles were published in the United States. Expanding on previous studies, Li, Xu, and Tian (2014) combined data and text mining techniques to provide enhanced fraud detection capability. The authors noted a lack of research on the combined use of the techniques. The researchers used a genetic algorithm to identify optimal parameters for the model, using financial and narrative data from 10-K filings. Li et al. (2014) found that the combined techniques increased the interpreting and explanatory power of the models.

While, (Thiruvadi & Patel, 2011) conducted a survey of data-mining techniques used in fraud detection and prevention and concludes that effective use of data mining techniques detect and prevent fraudulent activities and categorized four computer frauds where data mining tool can be employed: Management fraud; customer fraud; network fraud; and computer based fraud.

Then, (Gill & Gupta, 2009) researched on prevention and detection of financial statement fraud: a data mining approach and concludes that management fraud is a deliberate and wrongful act carried out of public companies using material misleading financial statement that cause damage to investors, creditors and the economic market.

Also (Kirkos, Spathis & Manolopoulos, 2007) conducted a research on data mining techniques for the detection of fraudulent financial statement. The study used a sample of 76 Greek manufacturing companies in order to inquire and draw an analogy between the performances of the various factors that are associated with the financial statements fraud. Neural networks Decision tree and Bayesian belief networks were the data mining techniques employed and the input data was the published financial statement contained falsified indicators; Bayesian belief networks performance was found out to be the best with 90.3% correct classification of the cross validation procedure, neural network had 80% success rate and decision tree model 73.6% success rate.

In another development, (Gill & Gupta, 2009) had a further study in which they used generic data mining framework for fraud prevention along with fraud risk-reduction for the financial statement fraud. The study divided data mining tasks into two groups of predictive tasks and descriptive tasks. Predictive data mining, along with machine learning helped in better fraud prevention, while performance evaluation of various data mining techniques using metrics such as error rate, information gain and Gini index for decision trees were employed.

Huan, Yan.,Yang and Hua (2008) conducted a research to assist auditors in identifying any possible fraud records and evaluating datasets by developing a fraud detection mechanism based on Zipf's law through simulation test and a case study. They used four key performance indicators, Audit Hit Rate, Bayers Audit Hit Rate, confusion matrix and the misclassification cost matrix. Finding showed that

ZipF's mechanism could be identified by ZipF's Analysis and this is more effective than a 100% sampling.

Kotsiantis, Kouthanakos, Tszolepis and Tompakos (2006), investigated the efficiency of the machine learning techniques in identifying firms that publish fraudulent financial statements. This they did by implementing a hybrid decision support system through combining algorithms that uses a stacking variant methodology. The data came from 164 non-financial Greek manufacturing firms listed, 41 of which had issued fraudulent financial statement. The study variables were collected from the financial statements of the firms. Results from this experiment indicated that the falsification indicators and a small list of ratios largely determined the classification result in published financial statements.

Liou (2006), investigated the similarities and differences between two models of fraudulent financial reporting detection and the business failure prediction that helped in identifying firms that procured losses. It aimed to find the effectiveness of the approach and the explanatory variables using data mining algorithms such as regression logistic, neural network, and classification trees to construct detection/prediction models using data from Taiwan Economic Journal data bank and Taiwan stock exchange corporation website. The financial variables were from 2003 to 2004. The findings show that the variables were significant in detecting fraudulent financial reporting and predicating business failures, logistic regression was considered the best of the three data mining algorithms.

Guo and Viktor (2008), researched on learning from skewed class multi-relation database. They focused the use of new strategy to address the imbalance in multi-relational data wherein one class in the target relation is higher than the others. The imbalances assist in diagnosing a disease or detecting a fraud case such as a credit-card fraud. Six benchmark data-sets were used for the experiment. The results indicated that imbalance in multi-relational method was better than other prevailing data mining algorithms in comparison, especially when there was a high class imbalance with regard to receiver operating characteristics curve and area under the curve.

Nonyelum and Chibueze (2009) employed the use of neural network technology and the rule-based components to develop credit-card fraud detection system using four clusters of low, high, risk and high risk using the two staged models that is frequently used in fraud detection. They developed a model identifying the behaviour of a cardholder and evaluating the transaction characteristic to detect fraudulent transactions using the self-organizing map algorithm. Other several models were generated by applying the artificial neural network trained with the unsupervised learning methods. This experiment further indicated that generation was done to secure a correct result and minimize the wrongful classification in which genuine transaction is considered fraudulent.

Xu, Sung and Liu (2007) also used data mining algorithm on simulated and real data to create user profile for identifying customer behaviour in detecting fraudulent transactions in an online system through a set of association rules. Anomalies were identified by comparing the incoming transaction of the user against that users profile based on his/her recent transaction. Conclusion is that the differences between the anomaly behaviour and the profiled user behaviour can be correctly interpreted by the proposed algorithm.

Graham & Patel (2006) also found in their studies that classification of network traffic helps to identity abnormal behaviours by detecting any derivations from the normal activity, (Kou., Peng., Chen & Shi, 2009) also examined network fraud and found out it is possible to use data mining-based network intrusion detection system and track the problem of solving the multi-class classification.

Becker, Volinsky and Wilks (2010) discussed different strategies and techniques used in the detection of the telecommunication-fraud history. They developed a fraud-management system to manage different types of fraud using call details, database required for storing data, fraud detection, algorithms fraud types and corrections and visualization cool that can help in diagnosis. (Liau et al, 2009) also examined the need for an effective and automated system for network forensic. The experiment results indicated that 91.59% of the attack types could be classified by the system thereby providing understandable information of forensic experts.

Sanver and Karahoca (2009) also compared the different data mining techniques, benchmarked each technique and identified Adaptive Neuro Fuzzy Inference for telecom-fraud detection in Turkey. The results showed that it provided 97% of sensitively, 99% of specificity, where 98.37 of the instances were correctly classified.

On the other hand (Koltler & Maloof, 2006) used machine learning and data mining to discover and classify malicious executable. The research selected executable which would appear undetected on a user's hard drive, without pre-processing or removing any obfuscation. The results showed that the boosted decision tree had an area under the ROC curve of 0.996, surpassing other models in fraud risk management.

Mukkamala, Sung & Abraham, (2005) showed that the ensemble of artificial neural network, SVM and Multivariate Adaptive Regression Splines, was superior to individual approach for intrusion detection in terms of classification accuracy. They used data from Massachusetts with five different classes of patterns. The results showed that 100% classification accuracies can be achieved if appropriate intelligent paradigms are chosen.

Another research on the use of business intelligence tools to detect fraud (Wang & Yang, 2009) found out that there is an increase in the use of data mining to detect fraud, but also lamented an overall underutilization.

Burnaby, Howe and Muehlman, (2013) made a review of the extent of the use of business intelligence to detect fraud. They came out with the following as regard the use of data mining: 15% use relational reporting; 13% use online analytical processing for fraud risk management; while other respondents complained that all the tools suggested were deficient and some noted that they use MS Access and the rest stated that they monitored email looking for transmission of credit card numbers.

Muhammed, Ghanbari and Einakian (2014) researched on using Data mining to detect fraud of internal audits by application of fraud deductive methods. The result shows that data analysis technology enables auditors and fraud examiners to analyse an organization's business data to gain insight into how well internal controls are operating and to identify transactions that indicates fraudulent activity or the highest risk of fraud.

More, (PWC, 2011) surveyed on Global Economic Crime, Cyber-crime (digital fraud) and reported that 45% indicated rising cybercrime fraud risks; 40% indicate that it is damaging reputation; 40% did not have capability to detect and prevent cyber-crime; 56% said the most serious fraud was an inside job and senior executives made up almost 50% who did not know if a fraud occurred and no indication of internal audit commitment in the use of data mining in cyber fraud risk management.

Lin, Chiu, Huang, and Yen (2015) examined the fraud triangle variables of fraud detection using data mining techniques. The authors provided a comparison to the results from a survey of experts. Lin et al. (2015) used the results to rank the fraud prediction factors by importance, providing a potential

solution to the budget and resource constraints of organizations for detecting fraud. The research included 129 fraud organizations and 447 non-fraud organizations and considered 32 factors for fraud assessment based on expert opinions. Lin et al. (2015) identified the top-ranking factors produced by the various data mining techniques as compared to expert decisions. The findings showed that two factors exist in the top 10 across all models examined. Those factors are the need for external financing and financial restatement frequency. The authors found the expert decisions to be consistent with the empirical results of the study and identified some gaps between the expert decisions and the prediction models. One advantage of the research is the use of objective measures. In addition, this research provided an effectiveness comparison of the detection tools used for the study. McMahon, Pence, Bressler, and Bressler (2016) contended that not every fraud can be identified using the fraud triangle methodology. The authors suggested adding a fourth element of capability, for example signing authority. McMahon et al. (2016) noted that it is difficult to detect fraud due to the deceptive nature of fraud, and that the top priority of management should be preventing fraud. Prevention can be implemented by using deterrence factors, such as surprise audits and whistle-blowing hotlines. This paper has indeed reviewed the related literatures on prevention and detection and the various techniques applied.

Methodology

The methodology of this research was a simple survey research design, i.e. primary data. The aim is to seek implementation opinion from banks control mechanisms how the explanatory variables of algorithms techniques on fraud detection in Nigeria banks.

Population of the study

The population of the study comprised all money deposit banks in Nigeria. The targeted population element of the study consist the three control mechanism (internal auditors, accountants and internal control) in each of the 15 banks of interest in this study with a population of 668.

Table 1: Population of the Study was made up of the internal audit staff, fraud audit staff and accountant staff in each of the 15 Banks used for the study

S/N	Description of Institution	Internal Audit Staff	Fraud Audit Staff	Accountants	Total
1.	Bank a	3	5	16	24
2.	Bank b	3	4	30	37
3.	Bank c	5	9	53	67
4.	Bank d	2	3	40	45
5.	Bank e	4	6	44	54
6.	Bank f	1	2	8	11
7.	Bank g	2	3	21	26
8.	Bank h	3	5	50	58
9.	Bank i	1	1	14	16
10.	Bank j	3	8	77	88
11.	Bank k	8	10	41	59
12.	Bank l	1	2	15	18
13	Bank m	6	7	58	71
14.	Bank n	2	3	76	81
15.	Bank o	1	2	10	13
		45	**70**	**553**	**668**

Source: Primary Survey Data sourced from the Banks by the Researcher.

Sampling Technique

The researcher employed Taro Yamane's formulae to determine the sample for the study. The formula is given as: $\boldsymbol{n = \frac{N}{1+N\,(e)^2}}$ where, n = Sample Size, N = Population Size (668)

E = Level of Significance (0.05), 1 = Constant. Using the formula, therefore, we have:

Sample Size $= \frac{\mathbf{668}}{\mathbf{1+668(0.05)^2}} = \frac{668}{669(0.0025)} = \frac{668}{1.67}$ = 400**,** distributed to the 15 banks.

Sources of Data

The Likert Scale response format of Strongly Agree (5 points), Agree (4 points), Strongly Disagree (3 points), Disagree (2 points), and Undecided (1 points), was used.

The researcher distributed and retrieved some of the questionnaire to the respondents while others were distributed through agents.

Methods of Data Collection, Analyses and Justification of the Statistical Tools Used.

Only questionnaires correctly filled and returned were used for analysis. Normal Likert values were 5, 4, 3, 2, and 1 and formula given as: $\overline{X} = \frac{\Sigma X}{n}$; where; $\overline{X}$ = Mean; X = the score; n = number of items and calculated as $\overline{X} = \frac{5+4+3+2+1}{5} = \frac{15}{5} = 3$. Our decision rule, therefore, is that any mean within 3.0 and above was considered as significant by the respondents, while a mean that is below 3.0 is taken as not significant. Descriptive statistics of percentage, mean and standard deviation were applied in the study.

To further strengthen the empirical analyses and test the posited hypotheses, ANOVA was employed to test the opinion of the three control mechanism and SPSS statistical analyses software was employed to carry out the analyses.

Reliability of the Instrument

The Cronback Alpha correlation of items calculated yield, 0.722, and is above the minimum stated by Cronback. (See appendix 2).

Data Presentation and Analysis

The return of 400 questionnaires distributed, showed that 213 (53%) were returned valid, while 169 (42%) were not returned and 18 (5%) were also returned but were invalid due to irregularities found in the responses.

Research Question One Analyses

Question I: To what extent do you agree that implementation of data mining algorithms-classification techniques detects fraud in banks?

Table 2 Responses on how Algorithms Classification Detects fraud in Banks

Job description	**To what extent do you agreed that algorithm classification detects fraud in banks**					
	Strongly disagree	Disagree	Undecided	Agree	Strongly agree	Total
Internal auditor	2 4.9%	1 2.4%	2 4.9%	17 41.5%	19 46.3%	41 100.0%
Internal control	0 0%	1 1.9%	7 13.5%	25 48.1%	19 36.5%	52 100.0%
Accountant	3 2.5%	8 6.7%	4 3.3%	48 40.0%	57 47.5%	120 100.0%

Source: Authors Computation, (2019)

Responses above show that, 2 (4.9) Internal Auditors IA and 0(0%) Internal Control IC and 3(2.5%) accountants strongly disagreed; 1(2.4%) of Internal Auditors IA, 1(1.9%) of Internal Control IC, 8(6.7%) of Accountants AC disagreed; 2(4.9%) of IA, 7(13.5) of IC, 4(3.3%) of AC were undecided; 17(41.5) of IA, 25(48.1) of IC, 48(40.0%) of AC Agreed; 19(46.3%) of IA, 19(36.5%) of IC, 57(47.5%) of AC strongly Agreed: that implementation of Algorithm Classification detects fraud in banks.

Question Two: To What Extent Do You Agree that the Implementation of Algorithms Clustering Techniques Detects Fraud In Nigerian Banks?

Table 3 Responses on How Algorithm Clustering Detects Fraud in banks

Job description	To what extent do you agreed that algorithms-clustering techniques detects fraud in banks					
	Strongly disagree	Disagree	Undecided	Agree	Strongly agree	Total
Internal auditor	1 2.4%	2 4.9%	6 14.6%	18 43.9%	14 34.1%	41 100.0%
Internal control	1 1.9%	2 3.8%	11 21.2%	27 51.9%	11 21.2%	52 100.0%
Accountant	3 2.5%	10 8.3%	18 15.0%	55 45.8%	34 28.5%	120 100.0%

Source: Authors Computation, (2019)

Respondents' opinion from the table above indicates that- On strongly agreed: IA scored 1(2.4); IC scored 1(1.9%), AC 3(2.5%); Disagree IA scored 2(4.9%), IC scored 2(3.8%), AC scored 10(8.3%); Undecided opinion: IA scored 6(14.6%), IC scored 11(21,2%), AC scored 18(15.0); while on Agree,

IA scored 18(43.9%), IC scored 27(51.9%), AC scored 55(45.8); and finally on Strongly Agree IA scored 14(34.1%), IC scored 11(21.2), AC scored 34(28.5%) that implementation of algorithms clustering detects fraud in banks.

Question Three: To What Extent Do You Agree that the Implementation of Algorithms Neural Networks Techniques Detects Fraud In Nigerian Banks?

Table 4: Responses on how Algorithms Neural Networks techniques detects fraud in banks

Job description	To what extent do you agreed that algorithms neural networks techniques detects fraud in banks					
	Strongly disagree	Disagree	Undecided	Agree	Strongly agree	Total
Internal auditor	1 2.4%	0 0%	5 12.2%	22 53.7%	13 31.7%	41 100.0%
Internal Control	1 1.9%	4 7.7%	5 9.6%	25 48.1%	17 32.7%	52 100.0%
Accountant	3 2.5%	4 3.3%	14 11.7%	59 49.2%	40 33.3%	120 100.0%

Source: Authors Computation, (2019).

The opinions above show that on Strongly Disagree IA scored 1(2.4%), IC scored 1(1.9%), AC scored 3(2.5%); Disagree IA scored 0(0%), IA scored 4(7.7%), AC scored 4(3.3%); on Undecided IA scored 5(12.2%), IC scored 5(9.6%), AC scored 14(11.7%); while on Agree IA scored 22(53.7%), IC scored 25(48.1%), AC scored 40(33.3%); and finally on Strongly Agree IA scored 13(31.7%), IC scored 17(32.7), AC scored 40(33.3%) that algorithms neural networks detects fraud in Nigeria banks.

Table 5: Mean and standard deviation scores on how data mining algorithms techniques detects fraud in banks

Variables	**Internal Auditors IA**			**Internal Control IC**			**Accountants AC**		
	Mean	Std	N	Mean	Std	N	Mean	Std	N
Algorithms classification	4.35	0.75	41	4.13	0.87	52	4.06	0.95	120
Algorithms clustering	4.30	0.91	41	3.85	1.30	52	4.85	0.60	120
Algorithms neural networks	4.12	0.94	41	3.97	1.19	52	4.93	0.10	120

Source: Authors Computation, (2019).

Remember that 3.00 is the mean score acceptance limit set for the study. From above AC, scored the highest mean (MN) of 4.93 and a low standard deviation (std) of 0.10 showing agreement of respondents. IC had the least mean of 3.85 and highest std of 1.30. Comparing the mean responses scored and the std, they are all acceptable. According to Uzoagulu, (1998); Albelson, (1985); Barz, (1963), when mean are so high and the std deviation so small it is an evident that they are tightly clustered around the mean and these show homogeneity and agreement among the respondents on the subject matter.

Testing of the Hypotheses of the Study

Table 6 ANOVA Hypotheses

		Sum of Squares	Df	Mean Square	F	Sig.
To what extent do you agree that algorithms classification techniques detects fraud in banks	Between Groups	5.101	4	1.275	0.04	0.97
	Within Groups	177.528	208	.854		
	Total	182.629	212			
To what extent do you agree that algorithms clustering techniques detects fraud in banks	Between Groups	8.705	4	2.176	0.37	0.69
	Within Groups	184.601	208	.888		
	Total	193.305	212			
To what extent do you agree that algorithms neural networks techniques detects fraud in banks	Between Groups	4.295	4	1.074	0.15	0.86
	Within Groups	165.648	208	.796		
	Total	169.944	212			

Source: Authors Computation, (2019)

SPSS ANOVA shows that all P values of the three explanatory variables of (0.97, 0.69, 0.86), were significant at 0.05 level limit of the study. The overall P value is 0.84. Applying the decision rule, the P value of 0.84 > 0.05, we therefore reject the null hypothesis and accept the alternate hypothesis that the implementations of all our explanatory variables of data mining algorithms techniques are significant in fraud detection.

Findings: The test result shows that all the null hypotheses were rejected, i.e. the alternative hypotheses were accepted. There were agreements that the implementation of data mining algorithms of: Classification; clustering; and neural networks techniques are significant in fraud detection in banks.

Discussion of Findings

Our study null hypotheses were rejected, while the alternatives were accepted. In other words, the respondents were united in both individual and group agreement on the subject matter. Any observed difference in their opinions was due to chance (Uzoagulu, 1989).

This study result does not in any way differ from the findings of some studies in other countries, (Alleyme, Persaud, Greenidge & Searly, 2010; Thiruved & Patel, 2011; Gill & Gupta, 2009; Kotsiantis, Kouthanakos, Tszolepis & Tompakos, 2006; Guo & Victor, 2008) who found that application of data mining detect and prevent fraudulent financial activities.

On the other hand, this result also differ from the conclusion of (Gill & Gupta, 2009), who applied data mining tool only on fraudulent financial reporting and (Liou, 2006) who also applied it on fraudulent financial reporting and loss on production and not generally on fraud generally.

More, this findings do not differ much in any form from the opinion of (Nonyelum & Chibueze, 2009; Uğurlu and Sevim, 2015), who found that data mining neural network detect fraud direction and (Xu, Sung & Liu, 2007), who also found out that data mining detect fraudulent transaction.

The findings do not differ from the findings of (Graham & Patel, 2006; Sanver & Karahova, 2009; Kolther & Maloof, 2006; Burnaby, Howe & Muehlman, 2013) who also agreed with the result that the application of data mining is significant in fraud detection.

Also, by the agreement of the respondents from the Likert analyses, shows that algorithms techniques of Classification, Clustering and Neural networks when implemented detects fraud in banks: is therefore not misleading, (Muhammed, Ghambari & Einakian, 2014; Burnaby, Howe & Muehlman, 2013; Wang & Yang, 2009).

If we discuss this finding considering the mean and standard deviation scores of the individual and group responses on the items within the data algorithms proxy, we will see that they were very high. Their mean scores were found to be above the study chosen limit of 3.0. Also, the standard deviations were reducing significantly to proof harmony in both the individual and group responses. However, considering the result of the Likert Scale analysis findings in all the items within the data algorithms proxy; showed that respondents had greater opinions of ‘‘strong and strongly agreed’’ opinions of responses.

Further from our literature reviews, we consider (Kirkos, Spathis & Manolopoulos, 2007), who concluded that data mining-neural networks had 80% success in detection of fraud; Uğurlu and Sevim, 2015), found neural network to be 90% success in fraud detection; while (Nonyelum & Chibueze, 2009), shows that the application of neural networks helps to detect fraudulent transaction and the wrongful classification in which genuine transaction is considered fraudulent.

Finally, (Mukkamala, Sung & Abraham, 2005), agreed that neural networks can be a good intrusion detection in terms of classification accuracy, while (Wang & Zaven, 2009; Patal & Zaven, 2010; Burnaby, Howe & Muechlmen, 2013; Muhammed, Ghanbari & Einakian, 2014; Sanver & Karahoca, 2009), all suggested that data algorithms detects fraud.

Summary of Findings

The study results revealed that all the null hypotheses tested were rejected and the alternate hypotheses were accepted. In others words, the respondents were of a strong opinion that: algorithms classification techniques; algorithms clustering techniques; and algorithms neural network are significant in fraud detection in banks.

Conclusions and Recommendations: The conclusions are anchored on the major findings: that the implementations of algorithms-Classification-Clustering and Neural networks techniques can detect fraud in banks. The recommendations are: Management and stakeholders of banks and other financial institutions that are susceptible to fraud are encouraged to implement these variable techniques in fraud detection. They should also carry out the training of their control staff in data mining techniques.

Current Research Contributions: The followings are considered as the current contributions of this study. **a)** This study found three variables and gave a construction of a new conceptual framework that showed how algorithms techniques can be implemented in fraud detection. **b)** The study also, has tried to bridge the gap of the paucity of academic studies and also provides a systematic empirical review of literatures covering data mining algorithms variables techniques in fraud detection which

may also serve as a reference to other researchers; and Finally, it provided a more understanding how to implement data mining algorithms in fraud detection through the suggestions of the research findings and recommendations.

Suggestions for Further Study: Further study can be carried out on the implementation of data mining algorithms techniques on fraud in other industries other than banks. Another study can also be carried effect of algorithms in fraud detection using secondary data in the analyses.

REFERENCES

Aggarwal, C.C. (2015). Data Mining. *Online* http://link.springer .com/openurlgenre=book&ISBN = DOI 10.1007/978-3-319-14142-8 10 285.

Ahmad, N., Othman, R., & Jusoff. (2009). The effectiveness of internal audit in Malasian public sector. *Journal of modern Accounting and Auditing,* 5(9), 784-790.

Albrecht, W.S., Albrecht, C.D., Albrecht, C.C. & Zimbelmain, M.F. (2011). Fraud examination. 4th Edn. *South Western Cengale Learning Manson, Ohw.*

Apostolou, B.A., Hassel, J.M., Webber, S.A. & Summers, G.E. (2001). The relative importance of management fraud risk factors. *Behavioural Research in Accounting* 13:1-24.

Araujo D.L.A., Lopes, H.S. & Freitas, A.A. (1999), "A parallel genetic algorithm for rule discovery in large databases", *Published in the Proc. Of IEEE Systems Man and Cybernetics Conference, Published By IEEE at Tokyo, V-3, pp 940-945*

Arun K.and Jabasheela L. (2014): "Big Data: Review, Classification and Analysis Survey", International Journal of Innovative Research in Information Security (ISSN: 2349-7009), V-1, I-3, pp 17-23

Companies and Allied Matter Decree CAMD, (1990). With the Amendment.

CPA. (2011). Employee fraud: A guide of reducing the risk of employee fraud and what to do after a fraud is detected. *Published by CPA Australia Ltd CAN 008392452.*

Cronback, L. J. (1960). Essentials of psychological testing. New York: *Harper and Brothers Publishers.*

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52, 281- 302.Retrieval from doi: 10.1037/h0040957 15/6/14.*

Duffield. G., Grabosky, P. (2001). The psychology of fraud. *Australian Institute of Criminology, Trends and Issues.*

Dunn, P. (2004). The impact of insider power on fraudulent financial reporting. *Journal of Management* 30(3), 397-412.

Dycus, D.E. (2002). Auditing for fraud. In Association of Certified Fraud Examiners Training Seminar. *Association of Certified Fraud Examiners. Available Fac...* (*Accessed on May 2014).*

Gbanbari, M.K. & Einakiam, M. (2014). Using data mining to detect frauds of internal audits. *Proceedings of 9th International Business and Social Science Research conference 6-8 January, Novotel World Trades Centre, Dubai. UAE.*

Gill, N.S. & Gupta, R. (2009). Prevention and detection of financial statement fraud: A data mining approach. *Journal of System Manager* (7), 55-68.

Graham, J., & Patel, S. (2006). Internet-Based security monitoring and control for utility companies and process plants: Data technology review. *International Journal of Business IT* (*3)*, 28-33.

Grazuoli, S., Jamal, K., & Johnson, P.E. (2006). Cognitive approach to fraud detection. *Journal of Forensic Accounting* (*7)*, 65-88.

Hamilton, D. I., Gabriel, J. M.O. (2012). Dimension of fraud in Nigeria quoted firms. *American Journal of Social and Management Sciences.*

Hinneburg, A., & Keim D. A (1999), "Optimal grid-clustering: Towards breaking the curse of dimensionality in high-dimensional clustering". Proceedings of the 25th International Conference on Very Large Data Bases held in the USA at 7-10 Sep pp 506–517.

Hodges, A. (2000). Emergency risk management. *Risk Management* 2(4), 7 – 18. *Published by Palamgrave Macmillan*.

Hoffman, V.B. (1997). Discussion of the effects of SAS no 82 on auditors' attention to fraud risk factors and audit planning decisions. *Journal of Accounting Research,* Vol. 35, 99-104.

Huang, S.M., Yen, D.C., Yang, L.W. & Hua, J.S. (2008). An investigation of zipf's law for fraud detection. *Decision Support System (46),* 70-83.

Igbataya, S. (2011). The challenges of the global economic, crisis and Nigerian's financial markets stability. *Journal of Emerging Trends in Economics and Management Sciences (JETEMS)* 2(6), 497-503.

Institute of Internal Auditors. (2009a). Global technology audit guide: Fraud prevention and detection in an automated world, *Altamonte Springs, FL*.

Jiang H, & Yang An. (2016): "Research on Pattern Analysis and Data Classification Methodology for Data Mining and Knowledge Discovery", *International Journal of Hybrid Information Technology (ISSN: 1738-9968), V-9, I-3, pp 179-188.*

Jiawei, H., Micheline, K & Pei Jian (2011):"Data Mining: Concepts and Techniques" *Published By Morgan Kaufmann, ISBN 978-9380931913.*

Kirkos, E.C., Spathis, C. & Manolopoulos, Y. (2007). Data mining techniques for the detection of fraudulent financial statement. *Expert System Application* (32), 995-1003.

Kolter, J.Z., & Maloof, M.A. (2006). Learning to detect and classify malicious exactness in the wild. *Journal of Machine Learning Research* (7*)*, 2712-2744. =

Kothari. C.K. (2004). Research methodology methods and techniques. *New Delhi, New Age Int. Publishers.*

Kotsiantis, S., Koumanakos, E., Tzehepis, D., Tampakas V. (2006). Forecasting fraudulent financial statement using data mining. *International Journal of Computational Intell.* (3*)*, 104-110.

Kou, G., Peng, X., Chen, Z. & Shi, Y. (2007). Multiple criteria mathematical programming for multi-class classification in network intrusion detection. *Information Society* (179), 371-381.

KPMG, L.L.P. (1994, 1995, 2013, 2014). Fraud survey. *Montvale, N.J; KPMG.* Online at www.sopac.org.au /Document-Library/fac. (Accessed on May 2014).

KPMG, LLP. (1994. 1993). Fraud survey. Montvale. Available Online at http:/www. Peterlang.com/ index.cfm? Event. Cst.ebook.datasheet & id = 21612 (Assessed may 2014).

KPMG, LLP. (1999, 1998). Fraud survey. *Montvale, N.J.* Available Online at http:/www. Peterlang.com/ index.cfm? Event. Cst.ebook.datasheet & id = 21612 (Assessed may 2014)

Lewis, D., Joseph, S.C., & Roach, K.Q. (2011). The implications of the current global financial economic crisis on integration: The Caribean experience.

Liao, N., Tain S. & Wang, T. (2009). Network forensics based on fuzzy logic and expert system. *Computer Communication* (32) 1991-1892.

Liou, F. M. (2006). Fraudulent financial reporting, detection and business failure prediction models: A comparism. *Management Auditing Journal* (23), 650-662.

Luehlfing, M.S., Daily, C.M., Philips, T.J., & Smith, L.M. (2003). Cyber-crimes intrusion, detection and computer forensic. *Internal Auditing* 18 (5), 9-13.

Manisha R. Thakare, & Mohod S.W. (2015): "Various Data-Mining Techniques for Big Data"Proceedings of the *International Conference on Quality Up-gradation in Engineering, Science and Technology Published by IJCA Held on Wardha India in 9-13 October 2015.*

McCue, C. (2015): Data mining and preventive analyses *(second edition)*

Muhammed, K., Ghambari, M. (2014). Using data mining to detect frauds of internal audit. *Proceedings of 9th international business and social sciences research conference Dubai UAE*

Mukkamala S., Sung, A.H., & Abraham, A. (2005). Intrusion detection using an ensemble of intelligent paradigms. *Journal of Network Computer Application* (28), 167-182.

Mukkamala S., Janoski, G. & Sung, A. (2002). Intrusion detection using neural networks and support vector machines. *Proceedings of IEEE International Joint Conference on Neural Network* pp. 1702-1707.

Murphy, K. R & Myors, B. (2003). Statistical power analysis: A sample and general model for traditional and modern hypothesis tests. *2nd Edition, Lawrence Eribean Associates*

Nunnally, J. C., Bernstein, I. H., & Berge, J. M. F. (1967). *Psychometric theory* (Vol. 2): New York, NY: McGraw-Hill.

Nonyelum, O.F., & Chibueze, I. H. (2009). Credit card fraud detection using artificial neural networks with a rule-based components *KFAI University Journal of Science Technology* (5), 40-47.

Nwana, O. C. (1981). Introductory to educational research. *Ibadan: Heinemann educational books ltd.*

O'Connell, J.J. (1977). Now tools for risk management research. *The Journal Issues and Practices* 1(1), 16-20 Online@ http://www.jstorrorg/stable/ 41942935./ Accessed April; 2014.

Orogun, W. (2009). Bank distress in history. *Nigeria, Burning Pot Com-Burning Port.Com.*

Osisioma, B. C. (2012). Combating fraud and white collar crime: Lessons from Nigeria. *2nd annual fraud & corruption African summit* 2012.

Owolabi, S.A. (2013). Fraud and fraudulent practices in Nigeria banking industry. *African Review 4*(3), 24-256.

Patel, S. & Zaveri, J. (2012). A Risk assessment model for cyber-attacks on information systems. *Journal of Computation* (5), 352-359.

Patil, D. V., & Bichkar, R.S. (2006), "A Hybrid Evolutionary Approach To Construct Optimal Decision Trees with Large Datasets", *Published in the Proc. Of IEEE International Conference on Industrial Technology, Published By in IEEE, Held in Mumbai, India on 15-17 Dec.*

Penton J. (2014). The role of data analytics in fraud prevention. Online at www.sopac.org.au /Document-Library/fac. (Accessed on May 2014).

PWC. (2011). Global economic crime survey (gecs) Online @ http://www.pwc.com/gx/en/weconomic-rime-survey/index. jhtm *5th June 2014l.*

Kurkos, E., Spathis, C. & Manolopoulos, Y. (2007). Data mining techniques for the detecting of fraudulent financial statements. *Expert System Application* (32), 995-1003.

Rammamoorti, S. (2008). The psychology and sociology of fraud: integrating the behavioural sciences component into fraud and forensic accounting curricula. *Issues in Accounting Education 23* (4), 521-533.

Sanver, M. & Karahoca, A. (2009). Using fraud detection: An adaptive neuro-fuzzy inference system in mobile telecommunication networks. *Journal of multiple-valued Logic Soft Compute* (155), 155-179.

Sherin, A., Uma S & Saranya K.K. (2014): "Survey on BIG Data Mining Platforms, Algorithm, and Challenges", *International Journal of Computer Science & Engineering Technology (ISSN: 2229-3345), V-5, I-9, pp 854-862*

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association, 69,* 730-737.

Stribu, D., Moraru, M., Farcane, N., Blidset, R., & Popa, A (2009). Fraud and error in auditors' responsibility levels. *Annales University Apulensis Series Occonomica*, 11(1), 5.

Thiruvadi, S., & Patel, S.C. (2011). Survey of data-mining techniques used in fraud detection and prevention. *Information Technology Journal*, (10) 710-716.

Uğurlu, M. & Sevim, S. (2015). Artificial neural network methodology in fraud risk prediction on financial statements: An empirical study in banking sector. *İşletme Araştırmaları Dergisi, Vol 7, Iss 1, Pp 60-89 (2015)*, (1), 60. Retrieved from 145 http://www.isarder.org/

United Nation. Conference on Trade and Development. 2010. Responding to the challenges posed by the global economic crisis to debt and development finance. *United Nations New York General.*

Uzoagulu, A.E. (1998). Practical guide to writing research project report in tertiary institutions. *Enugu, John Jacob's Classic Publishers Ltd.*

Wang, J. & Yang, J.G.S. (2009). Data mining techniques for auditing attestation function and fraud detection. *Journal of Forensic & Investigative Accounting* 1 (1).

Wolfe, J.B. (2012). Effective data mining for financial service companies. *The 11A Research foundation report. Third Quarter.* online@ http://www.the iia.org 20/5/1

XU, J., Sung, A.H., & Liu, Q. (2007). Behaviour mining for fraud detection. *Journal of Research and Practice Information Technology* (39), 3-18.