International Journal of African and Asian Studies
ISSN 2409-6938 An International Peer-reviewed Journal
Vol.31, 2017

Evaluation of the Contemporary Issues in Data Mining and Data Warehousing

Kalu, Alexanda Ogbonna Udu Bashir M. Ogungbangbe Nto Philips Chioma
Department of Marketing, College of Management Sciences
Michael Okpara University of Agriculture Umudike

Abstract
Over the past years data warehousing and data mining tools have evolved from research into a unique and popular business application class for decision support and business intelligence. This paper focuses on presenting the applications of data mining in the business environment. It contains a general overview of data mining, providing a definition of the concept, enumerating six primary data mining techniques and mentioning the main fields for which data mining can be applied. The paper also presents the main business areas which can benefit from the use of data mining tools, along with their use cases: retail, banking and insurance. Also the main commercially available data mining tools and their key features are presented within the paper. Theoretical and empirical literature was reviewed and various gaps in literature were identified. Besides the analysis of data mining and the business areas that can successfully apply it, the paper suggested and concluded that firms and scholars need to carry out more empirical research in the area of integrity of data mining and data warehousing since this will help eliminate marketing errors in operations and practice.

INTRODUCTION

Over the past decade, data mining became a matter of considerable importance due to the large amounts of data available in the applications belonging to various domains. Data mining is not static but a fast-expanding field, which applies advanced data analysis techniques, from statistics, machine learning, database systems or artificial intelligence, in order to discover relevant patterns, trends and relations contained within the data, information impossible to observe using other techniques (Berry & Linoff, 2011). The concept of Data Mining and Data Warehousing is gaining increasingly emphasis as a business information management tool that is expected to disclose knowledge structures that can guide decisions in conditions of limited certainty. A data warehouse guides business analysis and decision-making by creating an enterprise-wide integrated database of summarized, historical information. It integrates data from multiple, incompatible sources. By transforming data into meaningful information, and a data warehouse allows the manager to perform more substantive, accurate and consistent analysis (Agbonifoh, *et al.,* 2006; Pathak, *et al,* 2013).

In our contemporary society, companies collect huge volumes of data on a daily basis. Analyzing this data and discovering the meaningful information contained by it became an essential need for businesses. As the business environment develops and changes constantly, facing every day new challenges, the companies try to strengthen their market position and achieve competitive advantage by using new and innovative solutions, like data mining. Data mining solutions implement advanced data analysis techniques used by companies for discovering unexpected patterns extracted from vast amounts of data, patterns that offer relevant knowledge for predicting future outcomes.

A Data Base Management System is a collection of programs which provides the management of database. It has proper control access to the data and query language to retrieve the information. The database contains proper data up to the needs and occupies minimum storage space. The database contains no unnecessary data and we can added and update data.

A data warehouse is a collection of integrated databases designed to support a DSS. It is a collection of integrated, subject-oriented databases designed to support the DSS function, where each unit of data is non-volatile and relevant to some moment in time. Knowledge discovery (mining) in databases (KDD), knowledge extraction, data/pattern analysis, data archeology, data dredging, information harvesting, business intelligence, etc. Data mining is the process of finding patterns in a given data set. These patterns can often provide meaningful and insightful data to whoever is interested in that data. Data mining is used today in a wide variety of contexts- like in fraud detection, as an aid in marketing campaigns and even supermarkets use it to study their consumers. This is a perfect example of data mining – credit cards.

We can say that data warehousing is basically a process in which data from multiple sources /databases is combined into one comprehensive and easily accessible database. Then this data used by anyone who wants to mining the some useful data relates to field. A great example of data warehousing is social website that is facebook that everyone can relate to is what facebook does.it gathers all of your data –yours friends , likes ,comments and your stalk etc- and then stores data into the central repository. Data warehousing methodologies, as described by Zen *et al.,* (2005), share a common set of tasks, including business requirements analysis, data design, architectural design, implementation and deployment.

Virtually in every industry, progressive companies are moving toward the goal of understanding each customer individually and using that understanding to make it easier for the customer to do business with them rather than with competitors. These same firms are learning to look at the value of each customer so that they know which ones are worth investing money and effort to hold on to and which ones should be allowed to depart. This change in focus from broad market segments to individual customers requires changes throughout the enterprise, and nowhere more than in marketing, sales, and customer support.

Building a business around the customer relationship is a revolutionary change for most companies. Banks have traditionally focused on maintaining the spread between the rate they pay to bring money in and the rate they charge to lend money out. Telephone companies have concentrated on connecting calls through the network. Insurance companies have focused on processing claims and managing investments. It takes more than data mining to turn a product-focused organization into a customer-centric one. A data mining result that suggests offering a particular customer a widget instead of a gizmo will be ignored if the manager's bonus depends on the number of gizmos sold this quarter and not on the number of widgets (even if the latter are more profitable).

In the narrow sense, data mining is a collection of tools and techniques. It is one of several technologies required to support a customer-centric enterprise. In a broader sense, data mining is an attitude that business actions should be based on learning, that informed decisions are better than uninformed decisions, and that measuring results is beneficial to the business. Data mining is also a process and a methodology for applying the tools and techniques. For data mining to be effective, the other requirements for analytic CRM must also be in place. In order to form a learning relationship with its customers, a firm must be able to:

- Notice what its customers are doing
- Remember what it and its customers have done over time
- Learn from what it has remembered
- Act on what it has learned to make customers more profitable

Although the focus of this seminar is on the third bullet—learning from what has happened in the past—that learning cannot take place in a vacuum. There must be transaction processing systems to capture customer interactions, data warehouses to store historical customer behavior information, data mining to translate history into plans for future action, and a customer relationship strategy to put those plans into practice (Berry & Linoff ,2011).

MEANING OF THE CONCEPT

Data mining is the term used to describe the process of extracting value from a database. A data-warehouse is a location where information is stored. The type of data stored depends largely on the type of industry and the company. Many companies store every piece of data they have collected, while others are more ruthless in what they deem to be "important". Consider the following example of a financial institution failing to utilize their data-warehouse.

Another example of where this institution has failed to utilize its data-warehouse is in cross-selling insurance products (e.g. home, life and motor vehicle insurance). By using transaction information they may have the ability to determine if a customer is making payments to another insurance broker. This would enable the institution to select prospects for their insurance products.

These are simple examples of what could be achieved using data mining.

Four things are required to data-mine effectively: high-quality data, the "right" data, an adequate sample size and the right tool. There are many tools available to a data mining practitioner. These include decision trees, various types of regression and neural networks.

Income is a very important socio-economic indicator. If a bank knows a person's income, they can offer a higher credit card limit or determine if they are likely to want information on a home loan or managed investments. Even though this financial institution had the ability to determine a customer's income in two ways, from their credit card application, or through regular direct deposits into their bank account, they did not extract and utilize this information. Data mining is defined as a business process for exploring large amounts of data to discover meaningful patterns and rules (Pathak, *et al.,* 2013).

The data mining applications are available on all size systems for mainframe, client/server, and PC platforms. Data base mining or Data mining is a process that aims to use existing data to invent new facts and to uncover new relationships.

Data mining includes several steps: problem analysis, data extraction, data cleansing, rules development, output analysis and review. Data mining sources are typically flat files extracted from on-line sets of files, from data warehouses or other data source (Kurtz, 2008). Data may however be derived from almost any source. Whatever the source of data, data mining will often be an iterative process involving these steps. Following are the steps of data mining are:-.

1. Uniqueness Identification of the Objective -- Before it begins, one have to be clear on what you hope to

accomplish with your analysis. Know in advance the business goal of the data mining. Establish whether or not the goal is measurable.

2. Choice of the Data -- Once you have defined your goal, your next step is to select the data to meet this goal. This may be a subset of your data warehouse or a data mart that contains specific product information. It may be your customer information file. Segment the data as much as possible the scope of the data to be mined. Here are some key issues like

1.How current and relevant are the data to the business goal?

2. Are the data stable—will the mined attributes be the same after the analysis?

3. Compilation of the Data -- Once you've assembled the data, you must decide which attributes to convert into usable formats. Consider the input of domain experts/creators and users of the data. Establish strategies for handling missing data, extraneous noise, and outliers. Decide on a log or square transformation, if necessary. Determine the distribution frequencies of the data?

4. Evaluate the Data -- Evaluate the structure of your data. What is the nature and structure of the database? What is the overall condition and distribution of the dataset?

5. Choice of Appropriate Tools -- Two important factors for the selection of the appropriate data-mining tool for business objectives and data structure. Both should guide you to the same tool. No single tool is preferred to answer the queries

6. Prepare indented the Solution – Find out the answers of some questions like: What are the available format options? What is the goal of the solution? What do the end-users need graphs, reports, code?

7. Prepare the desired Model -- Now the data mining process begins. User split data into sets, construct and evaluate the model. The generation of classification rules, decision trees, clustering sub-groups, scores, code, weights and evaluation data/error rates takes place at this stage.

8. Check and Validate the Findings -- Share and discuss the results of the analysis with the business client or domain expert. Ensure that the findings are correct and appropriate to the business objectives. Find out the answers of many queries like-Do the findings make sense?

9. Reporting the Findings -- Prepare a final report for the business unit or client. The report should document the entire data mining process including data preparation, tools used, test results, source code, and rules. This report helps in decision making and plays important role in the growth of organization.

10. Combine components to integrate the solution – Share the findings with all interested end-users. You might wind up incorporating the results of the analysis into the company's business procedures. Although data mining tools automate database analysis, they can lead to faulty findings and erroneous conclusions if you're not careful.

NEURAL NETWORKS IN DATA MINING:

In more practical terms neural networks are non-linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data. Using neural networks as a tool, data warehousing firms are harvesting information from datasets in the process known as data mining. The difference between these data warehouses and ordinary databases is that there is actual manipulation and cross-fertilization of the data helping users makes more informed decisions. Neural networks essentially comprise three pieces: the architecture or model; the learning algorithm; and the activation functions (Berry & Linoff, 2011). Neural networks are programmed or trained to store, recognize, and associatively retrieve patterns or database entries; to solve combinatorial optimization problems; to filter noise from measurement data; to control ill-defined problems; in summary, to estimate sampled functions when we do not know the form of the functions. It is precisely these two abilities (pattern recognition and function estimation) which make artificial neural networks (ANN) so prevalent a utility in data mining. As data sets grow to massive sizes, the need for automated processing becomes clear. With their model-free estimators and their dual nature, neural networks serve data mining in a myriad of ways.

Data mining is the business of answering questions that you've not asked yet (Jiawei, *et al,* 2005). Data mining reaches deep into databases. Data mining tasks can be classified into two categories: Descriptive and predictive data mining. Descriptive data mining provides information to understand what is happening inside the data without a predetermined idea. Predictive data mining allows the user to submit records with unknown field values, and the system will guess the unknown values based on previous patterns discovered form the database.

Data mining models can be categorized according to the tasks they perform: Classification and Prediction, Clustering, Association Rules. Classification and prediction is a predictive model, but clustering and association rules are descriptive models. The most common action in data mining is classification. It recognizes patterns that describe the group to which an item belongs. It does this by examining existing items that already have been classified and inferring a set of rules. Similar to classification is clustering. The major difference being that no groups have been predefined. Prediction is the construction and use of a model to assess the class of an unlabeled object or to assess the value or value ranges of a given object is likely to have. The next application is

forecasting. This is different from predictions because it estimates the future value of continuous variables based on patterns within the data. Neural networks, depending on the architecture, provide associations, classifications, clusters, prediction and forecasting to the data mining industry.

Financial forecasting is of considerable practical interest. Due to neural networks can mine valuable information from a mass of history information and be efficiently used in financial areas, so the applications of neural networks to financial forecasting have been very popular over the last few years. Some researches show that neural networks performed better than conventional statistical approaches in financial forecasting and are an excellent data mining tool. In data warehouses, neural networks are just one of the tools used in data mining. ANNs are used to find patterns in the data and to infer rules from them (Kotler & Armstrong, 2010). Neural networks are useful in providing information on associations, classifications, clusters, and forecasting. The back propagation algorithm performs learning on a feed-forward neural network.

Feed forward Neural Network:

One of the simplest feed forward neural networks (FFNN), such as in Figure, consists of three layers: an input layer, hidden layer and output layer. In each layer there are one or more processing elements (PEs). PEs is meant to simulate the neurons in the brain and this is why they are often referred to as neurons or nodes. A PE receives inputs from either the outside world or the previous layer. There are connections between the PEs in each layer that have a weight (parameter) associated with them. This weight is adjusted during training. Information only travels in the forward direction through the network - there are no feedback loops.

Data Quality

Data quality is a key issue when an organization implements an enterprise wide data warehouse, for example, in customer relationship management (CRM) or other purposes. Utilizing CRM requires that customer information be of high quality in order to identify, validate and consolidate customers within an organization. Quality of the data will determine the quality of the data warehouse as well as the quality of the decision. In other words data quality is an investment in profitability (Sperley, 1999).

One main purpose for a DW creation is the possibility of having integrated data in one place (DW). It solves the problems with non-integrated data. But it does not really solve the problems with bad or incorrect data in the operational (source) systems. We may still suffer from the syndrome “Garbage in- Garbage out of Control” (Vaas, 2005).

Data mining and information analysis

Data mining is the use of mathematics, statistics, artificial intelligence and machine learning, to extract implicit, unknown and potentially useful knowledge from large, incomplete, noisy, fuzzy and random data. Data warehouse technology is a series of new independent application technology based on the needs of business information systems and the development of database systems. The essential difference between data mining and traditional data analysis methods (queries, reports, statistics and online analytical processing OLAP) is that data mining digs and finds that knowledge in the absence of clear information of assumptions. There are three characteristics in data mining model which are implicit, unpredictable, potential value.

Data mining, the extraction of hidden predictive information from large databases, is a powerful new technology with great potential to help companies focus on the most important information in their data warehouses. Data mining tools predict future trends and behaviors, allowing businesses to make proactive, knowledge driven decisions. Data mining tools can answer business questions that traditionally were time consuming to resolve. They scour databases for hidden patterns, finding predictive information that experts may miss because it lies outside their expectations (Pathak, *et al,* 2013).

Data warehouse and information collection

A data warehouse is of course a database, but it contains summarized information. Data warehouse refers to database that is maintained separately from an organizations operational databases. A warehouse holds read-only-data. Data mining, also called Knowledge-Discovery in Databases or Knowledge-Discovery (Pathak, *et al,* 2013). Data warehouse is a kind of data collection which is subject oriented, integrated, relatively stable, changing with time (different time) to support management in decision-making process, to help the data be geared to the needs of the subject, and to correspond to traditional database application. Subject will be a higher level of data classification standards for each subject area corresponds to a macro-analysis: integrated nature of the data warehouse is that before the data entering into the data warehouse, the data must be processed and integrated, which is a key step in the establishment of data warehouse. The contradictions in the raw data should be unified, but also the raw data should be changed from the application-oriented structure to the subject-oriented. Stability of the data warehouse is that data warehouse reflects the historical data, rather than data generated in daily transaction. After processed and integrated data entering into the data warehouse, there is little or no modification. Data warehouse is a collection of data at different times. It requires that retention of the data in data warehouse must meet the needs of decision analysis, and data must indicate the historical period of the data. With the data warehouse technology application deep going in recent years, data warehouse technology develops rapidly. Many companies already well accepted ‘the integration of data, find knowledge from the data,

use of data, and explanation with the data' and other new concept related to the improvement of all aspects of production, improvement of production efficiency and the development of the productive forces. Data are usually structured to be available in a form ready for analytical processing activities (e.g. online analytical processing (OLAP), data mining, querying, reporting and other decision supporting applications). A data warehouse is a subject-oriented, integrated, time-variant, non-volatile collection of data in support of management's decision-making process. (Efraim, T. *et al.,* 2010).

A data mart is a small data warehouse concentrated on a specific area of interest. Data warehouses can be subdivided into data marts for improved performance in use (Petre, 2013; Pathak, *et al,* 2013; Daskalaki, *et al,* 2003). The company can have one or more data marts towards a larger and more complex enterprise data warehouse. A Data Warehouse saves time of business user and helps to generate the reports quickly. Business users can quickly use these reports on one place and can take decisions quickly. Business users won't waste their precious time in collecting data from multiple sources. With the help of data warehousing, business can query the data themselves and saves money and time.

ONLINE ANALYTICAL PROCESS (OLAP)

After Data Warehouse has been populated with data, the next step in the data warehouse development is to provide the users with data analysis tool such as OLAP and Data Mining to analyze the business process for decision making (Akintola, K. G. *et al.,* 2011). OLAP is a service that automatically selects a set of summary views (tables), and saves these summary views to disk. OLAP also manage these views and update them when the fact table has new data. To create this, we must first create Analysis Services project using Business Intelligence Development Studio. Then, we define the data source, create the data source view and then create Cube. Figure below shows a screen shot of the OLAP cube created with SQL Server.

Data Mining Applications in Marketing and Customer Relationship Management

Some people find data mining techniques interesting from a technical perspective.

However, for most people, the techniques are interesting as a means to an end. The techniques do not exist in a vacuum; they exist in a business context. This paper is about the business context. There is a set of business objectives that can be addressed by data mining and data warehousing (Michael, 2009). Each of the business objectives is linked to specific data mining techniques appropriate for addressing their problems. This can be seen in complexity of the customer relationship, the problem of communicating with potential customers about whom little is known, and works up to the varied data mining opportunities presented by ongoing customer relationships that may involve multiple products, multiple communications channels, and increasingly individualized interactions.

Prospecting seems an excellent place to begin a discussion of business applications of data mining and warehousing. After all, the primary definition of the verb to prospect comes from traditional mining, where it means to explore for mineral deposits or oil. As a noun, a prospect is something with possibilities, evoking images of oil fields to be pumped and mineral deposits to be mined. In marketing, a prospect is someone who might reasonably be expected to become a customer if approached in the right way. Both noun and verb resonate with the idea of using data mining to achieve the business goal of locating people who will be valuable customers in the future.

For most businesses, relatively few of Earth's more than six billion people are actually prospects. Most can be excluded based on geography, age, ability to pay, and need for the product or service. For example, a bank offering home equity lines of credit would naturally restrict a mailing offering this type of loan to homeowners who reside in jurisdictions where the bank is licensed to operate. A company selling backyard swing sets would like to send its catalog to households with children at addresses that seem likely to have backyards. A magazine wants to target people who read the appropriate language and will be of interest to its advertisers, and so on.

Data mining can play many roles in prospecting. The most important of these are:

- Identifying good prospects
- Choosing a communication channel for reaching prospects
- Picking appropriate messages for different groups of prospects.

Although all of these are important, the first—identifying good prospects— are the most widely implemented (Berry & Linoff ,2004). Data mining and warehousing can be used to improve direct marketing campaigns, Airways ticket and flight control, Provide decision tree for informed decision budget, strategically position a form for sustainable profitability and competitive advantage, Customer Relationship management, segmenting the customer base, marketing research and product development, reducing exposure to credit risk, determining customer value, customer loyalty and retention and manufacturing process control.

DATA MINING SOLUTION FOR THE BUSINESS ENVIRONMENT

The main features of a data mining solution for the business environment, includes as follows;

- **Selecting the data**: identify the data sets used for a specific analysis and improve the initially selected data sets if required;
- **Preparing the data:** transform and clean the data so it is in the appropriate format for applying data mining techniques;
- **Choosing the data mining technique:** select the algorithm associated to a data mining technique that is suitable for the required analysis;
- **Configuring the settings for the data mining technique**: configure for the selected algorithm the necessary parameters;
- **Executing the data mining process:** execute the configured data mining process;
- **Viewing the results of the data mining process:** view the results generated by the execution of the data mining process.

The data is extracted from the data sources, both operational systems and data files, and loaded through the ETL (Extract, Transform and Load) process to the data warehouse. The data warehouse can contain raw data – the data in a detailed format, as it has been extracted, summary data – data that has been aggregated and transformed – and metadata, data that provides information about the raw and summary data belonging to the data warehouse (Petre, 2013). The data belonging to the data warehouse can be organized in data marts.

The solution provides a data mining engine that may be used for obtaining advanced analysis. The solution has a graphical user interface that provides access to the main features of the solution, as presented in the article: select and prepare the data, choose data mining technique, configure the settings and execute the data mining process, view the results obtained.

The business user may access the GUI of the solution from various devices, like the desktop, laptop, tablet or mobile. Using the functionalities described above the solution allows accessing and analyzing business related information in order to obtain valuable knowledge concerning the business.

DATA WAREHOUSE AND DATA MINING: APPLICATION AREAS IN BUSINESS

Data warehousing and Data Mining has gained improved popularity in multiple areas of business to analyze the large databases quickly which would be too complex and time consuming. Some of these application areas are listed below.

1) Government: for searching terrorist profile and threat assessments.
2) Finance: analysis and forecasting of business performance, for stock and bond analysis.
3) Banking: to learn underwriting, mortgage approval etc.
4) Direct marketing: for identifying prospects that are included in mailing list so as to obtain highest response time.
5) Medicine: for drug analysis, diagnosis, quality control and epidemiological studies.
6) Manufacturing: for improved quality control and maintenance.
7) Churn analysis: to predict customers who are likely to quit the company and move to a competitor company.
8) Market segmentation: to identify customer's common characteristics and behavior that purchases the same products of a company [20].
9) Trend analysis: to analyze the difference between the customer's behavior over consecutive months.
10) Fraud detection: to identify the fraud users in telecommunication industry as well as credit card usage.
11) Web marketing: for advertisements and personalization opportunities (Kotler & Armstrong, 2010; Kurtz, 2008).

ROLE OF DATA WAREHOUSING AND DATA MINING IN DECISION

The goal of a data warehouse is to support decision making with data. Data mining can be used in conjunction with a data warehouse to help with certain types of decisions (Berry & Linoff, 1997; Hui & Jha, 200; Pathak, *et al,* 2013; Ling & Li, 1998). To be successful, data is warehousing and data mining needs a skilled user who will supply the correct data and a specialist who can make objective conclusions out of the output that is created. If the user supplies incorrect or minimal amount of information, output will be affected and forecast will not be credible.

Data warehousing and data mining plays an important role in decision making of the organization. Data warehousing provide answers of many queries to the organization and the user and helps in decision making. There are many types of queries of the organization like tactical query, strategic query, and update query.

A tactical query is a database operation that attempts to determine the best course of action right now (Berry & Linoff, 2004). Whereas the strategic query provides the information necessary to make long term business decision, a tactical query provides information to rank and file elements in the field that need to respond quickly to a set of unfolding events. Tactical queries tend to produce a very small result set. It is not uncommon for the result set to be less than a dozen rows. Usually the result set is designed to fit into a single window on a display screen.

A strategic query is a database operation that attempts to determine what has happened, why it happened, and/or what will happen next. It typically accesses vast amounts of detailed data from the warehouse and ranges in complexity from simple table scans to multi-way joins and sub queries.

Applications that generate strategic queries include; report generation, OLAP, decision support, ad-hoc, data mining, etc. An update query is a database operation that modifies the state of a database. Teradata provides a set of bulk load utilities used to load large quantities of data into the database in an efficient fashion.

EMPIRICIAL REVIEWS

Peng, *et al.,* (2011) investigated the application of business intelligence technology in Airports and Airlines companies and concluded that with the continuous development of airports and airlines, airports and airlines will continue to increase in volume of business. That the amount of various data of airports and airlines becomes very large, large amount of data can affect the speed of the decision maker. This implies that the Business intelligence technology applied to airports and airlines will assist decision-makers to make more timely and effective, more scientific and reasonable decision-making. Decision-makers' decision-making will be on a higher level. This can also be seen as a source of decision tree in strategic marketing planning which helps the organization to make informed decisions that will translate to sustainable profitability and business growth.

Petre, (2013), concluded that the architecture proposed for the data mining solution for the business environment would improve the efficiency of a company, by providing valuable decision-making knowledge to minimize operating costs and gain competitive advantage. This is true because over the years rapid technological changes and applications to business have changed the way businesses are operated in our contemporary society. Banking service have really improved with a great measure because of data mining, one can deposit his money in Nigeria and make withdrawals from financial institution in USA. Also in the area of marketing research and development, firms have access to data about their target market and this help them to produce products that will satisfy their needs and also help them in gaining competitive advantage.

Pathal, *et al.,* (2013), concluded that Data Warehouses and Data Mining can be used in organizations and that it helps in decision making, and allows the manager to perform more accurate, substantive and consistent analysis. This is evident in the smooth running of the organizational operations in the most effective and efficient manner. Notwithstanding the huge volume of data that is being handled for optimization of their operations. Data warehouse is a database used for reporting and analysis. It is a place where data is stored by integrating different Data bases. It can be used for storing current and historical data. With the help of historical and current database new prediction can be drawn.

Both domestic and foreign experts have done some research about data warehouse applications in airports and airlines. Harry K.H. Chow, King Lun Choy, W.B. Lee, K.C. Lau (2006) proposed the management of the airport's cargo through the application of RFID-based resource management system. By using Radio frequency identification, case-based reasoning technology to improve resource management efficiency, the practice prove that using data warehouse technology can improve the management efficiency, save time and reduce the cost. This have really improved on the airways logistical activities and also digitalized the cargo handling processes. Never the less, some constraints are still being faced in handling cargos, which is as a result of inconsistence of data.

T. C. Poon, K. L. Choy, Harry K. H. Chowa, Henry C. W. Lau, Felix T.S. Chan, K. C. Ho (2009) used radio frequency identification technology based on logistics and resource management system to manage data. Radio frequency identification technology is used to get the airport cargo data to test the performance of RF equipment. Three objectives are achieved: (i) a simplification of RFID adoption procedure, (ii) an improvement in the visibility of warehouse operations and (iii) an enhancement of the productivity of the warehouse. At the end of the article, author gave the technology high evaluation and he thinks that the technology is available in practice. S G Li, X. Kuo (2008) proposed an enhanced fuzzy neural network (EFNN) technology based on the central warehouse management reserve inventory decision support system. Traditional neural networks are found to be suffered from the problem of low accuracy of forecasting unseen examples. Therefore, in this EFNN, the following improvement is made: First, it assigns connection weights based on the fuzzy analytic hierarchy process (AHP) method without painstakingly turning them. Second, by generating and refining activation functions according to genetic algorithm, EFNN can provide comprehensive and accurate activation functions and fit a wider range of nonlinear models. Last, but not least, an adaptive input variable is introduced to decrease the impact of the bullwhip effect on the forecasting accuracy.

Shijin Wang, Dong Sui, Bin Hu (2010) used gray mixed model to predict air flow. With the historical data analysis of the national-wide air traffic from the year 1985 to 2008, the forecasting applicability is examined with three different methods, such as the time series forecasting, regression forecasting and neural-network forecasting. The result shown that predict accuracy of gray mixed model was better than the other three models.

PERCEIVED GAPS IN LITERATURE

From the review so far, it was found that these data mining and data warehouse technological advances have not been simplified to meet the compatibility business environment of small and medium scale Enterprises in developing countries and as such, only multinationals are basically reaping the benefit of its technological innovations and the benefit thereof. As such, the system should be less complex so that that average business outfit can teach it staff on its usage and applicability for the smooth operation of their business and also in making informed decisions as this will go a long way in helping them to achieve sustainable competitive advantage.

DW and DM is not an end in itself. The knowledge gained and business intelligence attained by DW and DM could be retained and reused by using CBR (Case-based Reasoning) technology. It is the technology for solving new problems by adapting the known solutions of previous similar problems (Chan and Chowdhury, 2005). Data mining and CBR can be seen as complementary methodologies for knowledge management in a broader sense. Today's reporting environment give users access their data, but it does not solve all the problems of user. The users have privilege to access the data but do not guarantee the integrity of the data and adequacy of response time. There are no studies that have extensive addressed on how the integrity of the data drawn from the data ware houses can be ascertained.

Data mining and data warehousing system in the world at large have failed in decisions taken from the database transforms to marketing error. This is mostly in the developing nations and emerging economies of the world. Firms and scholars need to carry out more empirical research in the area of integrity of data mining and data warehousing since this will help eliminate marketing errors in operations and practice.

SUGGESTIONS

From the analysis so far, and in order to curtail on poor decision outcomes, firms should try to verify the data from the data warehouse before mining and onward usage for decision making, as this will go a long way in limiting the occurrences of poor marketing outing and marketing errors.

Numerous roles and responsibilities will need to be acceded to in order to make data warehouse efforts successful and generate return on investment. For the technical personnel (application programmer, system administrator, database administrator, data administrator), it is recommended that the following roles be performed full-time by dedicated personnel as much as possible and that each responsible person receive specific Data Warehouse training. The data warehouse team needs to lead the organization into assuming their roles and thereby bringing about a partnership with the business. Management also needs to make actionable plans out of these directives and make sure the staff executes on them.

Business is well-fit domain for applying data mining as it provides large volumes of data. In a business context, the successful introduction of data mining requires using data mining techniques to address a real business challenge. For companies that are just getting started with analytical customer relationship management, integrating data mining can be a daunting task. A proof-of-concept project is a good way to get started. The proof of concept should create a solid business case for further integration of data mining into the company's marketing, sales, and customer-support operations. This means that the project should be in an area where it is easy to link improved understanding gained through data mining with improved profitability.

There is scarcity of literature in empirical review of data mining and data warehousing in business operations and as such academic scholars and organizations need to carry out more extensive empirical studies in order to build up knowledge in this area.

CONCLUSION

Data mining is used in support of both advertising and direct marketing to identify the right audience, choose the best communications channels, and pick the most appropriate messages. Prospective customers can be compared to a profile of the intended audience and given a fitness score. Should information on individual prospects not be available, the same method can be used to assign fitness scores to geographic neighborhoods using data of the type available from the Nigeria Bureau of Statistics, Nigeria Statistical Bulletin, U.S. census bureau, Statistics Canada, and similar official sources in many countries.

Information about current customers can be used to identify likely prospects by finding predictors of desired outcomes in the information that was known about current customers before they became customers. This sort of analysis is valuable for selecting acquisition channels and contact strategies as well as for screening prospect lists. Companies can increase the value of their customer data by beginning to track customers from their first response, even before they become customers, and gathering and storing additional information when customers are acquired.

Once customers have been acquired, the focus shifts to customer relationship management. The data available for active customers is richer than that available for prospects and, because it is behavioral in nature rather than simply geographic and demographic, it is more predictive. Data mining is used to identify additional

products and services that should be offered to customers based on their current usage patterns. It can also suggest the best time to make a cross-sell or up-sell offer. One of the goals of a customer relationship management program is to retain valuable customers. Data mining can help identify which customers are the most valuable and evaluate the risk of voluntary or involuntary churn associated with each customer. Armed with this information, companies can target retention offers at customers who are both valuable and at risk, and take steps to protect themselves from customers who are likely to default. Data mining comes in two forms. Directed data mining involves searching through historical records to find patterns that explain a particular outcome. Directed data mining includes the tasks of classification, estimation, prediction, and profiling. Undirected data mining searches through the same records for interesting patterns. It includes the tasks of clustering, finding association rules, and description. Data mining brings the business closer to data.

Data warehousing solve the above problems and provide technology which enables the user or decision maker to process the huge amount of data in a short amount of time. With the help of data warehousing, user extract the knowledge in a real time and its help the user in the decision making. Many companies want to use that data for other purposes. So data mining techniques are evolved for extracting new knowledge from data warehouse.

Data warehousing and data mining provide the right foundation for building decision support and executive information system tools which help to measure the progressing speed of organization toward its goal. Data warehousing and data mining provide a technology that enables the user or decision-maker in the corporate sector/govt. to process the huge amount of data and make decisions which are useful for whole organization.

Data warehousing and Data Mining are now playing a significant role in strategic decision making. It helps companies make better decisions, streamline work-flows, and provide better customer services. The power of Data-Warehousing in data analysis is tremendous and data-mining can discover hidden treasures in the data-warehouse. Data Warehouse and Data Mining technologies have bright future in business applications as it helps to generate new possibilities by automated prediction of trends and behaviors in a large database. Data mining techniques help to automatically discover the unknown patterns like identifying anomalous data that highlight errors generated during the data entry.

Data Warehouse and Data Mining technologies have become a hit with various industries like sales & marketing, healthcare organization, financial institutions and many more. These technologies have a lot of benefits in varying fields. Data Warehouse & Data Mining are very essential components in business operations to gain competitive intelligence. These technologies allow statistical multidimensional analysis of data to evaluate relationships, correlations and trends in business.

LIMITATIONS OF THE STUDY/SUGGESTION FOR FURTHER STUDIES

The study's limitation is based on the fact that this paper is purely a theoretical paper and no empirical evidence was incorporated. As such it is therefore suggested that future studies should include empirical data in their study, the study should be focused on a particular industry.

REFEFRENCES

Akintola K.G., Adetunmbi A.O. and Adeola O.S. (2011). Building Data Warehousing and Data Mining from Course Management Systems:: A Case Study of Federal University of Technology (FUTA) Course Management Information Systems. Nigeria Computer Society (NCS): 10th International Conference – July 25-29

Berry, M.J.A. and Linoff G.S. (2004). Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management (paperback), 2nd Edition. Canada: Wiley Publishing Inc.

Berry, M.J.A. and Linoff G.S. (2011). Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management (paperback),. 3rd Edition. USA: Wiley Publishing Inc.

Berry, Michael J.A., and Gordon Linoff (1997)," Data mining techniques: for marketing, sales, and customer support", New York: Wiley.

Chowdhury, S.(2003). Databases, Data Mining and Beyond. *The Journal of American Academy of Business, Cambridge. Vol. 2, No. 2, pp. 576-580.*

Chan, J and Chowdhury, S. (2005). Enhancing Customer Services using Case based Reasoning. *The communications of the ICISA – the official journal of the International Chinese Information Systems Association,* 7(1), 53-61, Summer.

Chowdhury, S. (2005). A Holistic Framework for Knowledge Management. In Issues in Information Systems. *International Association for Computer Information Systems (IACIS).* 6(2), 10-16.

Devale, A. D. and Kulkarni, R. V. *(2012).* Applications of data mining techniques in life insurance. *International Journal of Data Mining & Knowledge Management Process, 2(4), 31-40, July.*

Daskalaki, S., Kopanas, I., Goudara, M., and Avouris, N. (2003). "Data mining for decision support on customer insolvency in telecommunications business", *Europeon Journal of Operational Research,* 145(2), 239-

255.
Efraim T., Jay E., Teng-Peng L., Ramesh S, (2010). Decision Support and Business Intelligence Systems (8th edition). USA: Prentice Hall.
Hui, S. G. Jha(2000), Data Mining for Customer Service Support, Information & Management, *Elsevier*. 12(4), 123-127.
Harry K.H. Chow, King Lun Choy, W.B. Lee, K.C. Lau. (2006). Design of a RFID case-based resource management system for warehouse operations. Expert Systems with Applications, 30(4), 561-576, May.
Haykin, S. (1999). Neural Networks. Prentice Hall International Inc., 1999
Jiawei Han, Micheline Kamber, Jian Pei(2005), Data Mining: Concepts and Techniques, 2nd edition, Morgan Kaufmann.
Jiawei, H., Micheline, K. and Jian P. (2011). Data Mining: Concepts and Techniques. 3rd Edition. USA: Morgan Kaufmann Publishing.
Kotler, P. and Armstrong G. (2010). Principles of Marketing. New Jersey: Pearson Prentice Hall.
Kurtz, D.L. (2008). Contemporary Marketing. (13th Edition). New York: South Western Cengage Learning.
Ling, C. X. and Li, C.(1998). Data mining for direct marketing: Problems and solutions, in Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining.
Li, X. G. and Kuo, X. (2008). The inventory management system for automobile spare parts in a central warehouse. Expert Systems with Applications, 34(2), 1144-1153, February.
Lungu, I. and Bâra, A. (2012). Improving Decision Support Systems with Data Mining Techniques, "Advances in Data Mining Knowledge Discovery and Applications" – chapter 18, Croatia: InTech Publisher.
Michael A. King (2009). A Realistic Data Warehouse Project: An Integration of Microsoft Access and Microsoft Excel Advanced Features and Skills Journal of Information Technology Education 8(1), 231-240.
Oded Maimon and Lior Rokach (2010). Data Mining and Knowledge Discovery Handbook. 2nd Edition. USA: Springer Publishing Company.
Pathak, M., Singh, S. and Oberoi, S. S. (2013). Impact of Data Warehousing and Data Mining in decision making. *International Journal of Computer Science and information Technologies*. 4(6), 995-999.
Petre, P. (2013). Data Mining Solutions for the Business Environment. *Database Systems Journal*, 4(4), 21-29.
T.C. Poon, K.L. Choy, Harry K.H. Chow, Henry C.W. Lau, Felix T.S. Chan, K.C. Ho. (2009) A RFID case-based logistics resource management system for managing order-picking operations in warehouses. Expert Systems with Applications, 36(4), 8277-8301, May.
Vikas Jayasree and Rethnamoney Vijayalakshmi Siva Balan (2013), A Review on Data Mining in Banking Sector, *American Journal of Applied Sciences*, 10(10), 1160-1165.