ISSN 2455-4863 (Online) **www.ijisset.org** *Volume: 6 Issue: 1 | 2020*

The Imperatives of Data Mining in Business and Management Sciences

Frank Alaba Ogedengbe (PhD) [1], Amal Abdulghani Altalhi [2]

[1] *Department of Business Administration/Management, Auchi Polytechnic, Nigeria*

[2]*Department of Management information Systems, College of Business Administration, University of Tabuk, Saudi Arabia*

Abstract: *The imperatives of data mining in business and management science highlight the importance, relevance and role of data mining in business and management science. The literature method was employed to painstakingly highlight the core issues in this article which is the imperatives of data mining. The study looks at the concept, techniques of discovering interesting patterns, data mining applications and the imperatives of data mining in business and management sciences. How data mining can be employed to combat crime, insecurity and other related vices in the society were equally examined. Some drawbacks and challenging issues capable of impeding the enormous benefits of data mining were discussed. Based on the enormous benefits and advantages of data mining, we conclude that attention be given to data mining in management science, business operations and governance.*

Keywords: *Data Mining, Business and Management Sciences.*

INTRODUCTION

Technology plays a major and remarkable role in business, communication, management and national development, and it needs to be said that technology runs on data. Technology has proven very useful in sundry fields like government, institutions' management and business, as well as other varying fields of human endeavour. This is evident in the seamless way and manner information dissemination has become very fast and almost instantaneous in every sense of it. This is so much so that whatever is happening in a particular place right now is almost known everywhere in the world at the same time. For example, news now travels as fast as almost the speed of light, moving from one end of the globe to the other very quickly and seamlessly. These are made possible by the use and systematic handling of data in varying forms, degrees and dimensions. Data are the ingredients technology uses in information gathering, processing and dissemination.

Data are becoming very large in every way and in every sense of it. Individual and personal data, community and public data, as well as corporate and institutional data, are growing in size, shape and dimension. This requires that we equip ourselves to be able to sift the relevant and required data from the mass of data available which come on a very large and almost unmanageable magnitude. This is required on a regular and continuous basis. The knowledge of data mining becomes pertinent in this regard.

Data mining is the process of discovering interesting patterns and knowledge from large amount of data. It is the process of extracting useable data from a large set of raw data.

It is the process of analyzing data patterns in large batches of data using one or more approaches or software. Data mining is also known as Knowledge Discovery in Data (KDD). These data come from various data sources which include databases, data warehouses, the Web, other information repositories, or data that are streamed into the system dynamically. Data mining is a process of turning raw data into useful information. It uses software to look for patterns in large batches and mass of data. It is a process of discovering hidden valuable knowledge by analysing large amounts of data, which is stored in databases or data warehouses, using various data mining techniques such as machine learning, artificial intelligence and statistical techniques (Rajkumar, 2014).

Business is the activity of making one's living or making money by producing or buying and selling products, such as goods and services. Put in another form, business is any activity or enterprise entered into for the purpose of making profit. A business is defined as an organisation or enterprising entity engaged in commercial, industrial or professional activities. The term business also refers to the organized efforts and activities of individuals to produce and sell goods and services for profit maximization. A business is an organisation or economic system where goods and services are exchanged for one another or for money.

A business may be set as Sole Proprietorship, Partnership and Corporation. A sole proprietorship is a business owned and operated by a single natural person. There is no legal separation between the business and the owner; the tax and legal liabilities of the business are thus that of the owner. A partnership is a business relationship between two or more people who join to conduct business. Each partner contributes resources and money to the business and shares in the profits and losses of the business. The shared profits and losses are recorded on each partner's tax return. A

corporation is a business in which a group of people acts together as a single entity; most commonly, owners of a corporation are shareholders who exchange consideration for the corporation's common stock. All businesses must address six basic activities, Human Resources, Customer Service, Accounting, Budgets, Sales and Marketing (Sainz, 2019).

Management science is the broad interdisciplinary study of problem solving and decision making in human organisations, with strong links to management, economics, business, engineering, management consulting, and other fields. Management science helps businesses to achieve goals using various scientific methods. Management science theory is management that focuses on the use of techniques to maximize the use of organisational resources to produce goods and services. It is a contemporary approach to management that is an extension of scientific management that measures the worker to task mix and ratio to raise efficiency. Management science generally refers to mathematical or quantitative methods for business decision making. The breadth of management science techniques includes; mathematical programming, linear programming, simplex method, dynamic programming, goal programming, integer programming, nonlinear programming and stochastic programming, to mention just a few (Ogedengbe, Ogedengbe & Ovuworie, 2018).

The imperatives of data mining in business and management sciences highlight the importance, relevance and role of data mining in business and management science. This paper looks at the concept, techniques of discovering interesting patterns, data mining applications and the imperatives of data mining in business and management science. How data mining can be employed to enhance efficient and effective decision making in business, management and other areas like combating crime, insecurity and other related vices in the society are explained. The drawbacks and issues capable of impeding the enormous benefits of data mining are also discussed.

DATA MINING

Data mining is the process of exploring and uncovering patterns in large data sets. Large data set is also called Big Data. It is a subfield of computer science which blends many techniques from statistics, data science, database theory and machine learning. It is the process of discovering patterns in large data set involving methods at the intersection of machine learning, statistics and database systems. Data mining is an essential process where intelligent methods are applied to extract useful data from a large set of data and also to determine useful data patterns (Rajkumar, 2014).

Data mining is commonly defined as the computer assisted search for interesting patterns and relations in large databases. It is an area of research that builds on the older disciplines of statistics, databases, artificial intelligence or machine learning and data visualization. The emergence of data mining is often explained by the ever increasing size of databases together with the availability of computing power and algorithms to analyse them. Data mining is usually considered to be a form of secondary data analysis. This means that it is often performed on data collected and stored for a different purpose than analysis. They are usually data for administrative and decision making purposes (Vel, Anderson, Coney & Mohay, 2001).

The overall goal of data mining is to extract information from a data set, and transform it into an understandable structure and usable form for decision making in business, management, government and other human fields (Yamuna & Bhuvaneswari, 2012).

DATA MINING PROCESS

Data mining is digging through mountain of data to extract useful information and looking for sense among the mass of raw data. It is the computer assisted process of digging through and analyzing enormous sets of data and then extracting the meaning of the data. It is a process used by companies to turn raw data into useful information. Data mining is divided into two parts called data preparation and data mining. The Data preparation process includes data cleaning, data integration, data selection and data transformation. Data mining include, pattern evaluation, knowledge representation and integration (Brown, 2014; Saltos & Cocea, 2017). Data mining process includes the following steps;

Identify information source: One needs to identify the data and data source as well as the sources of information, and from that be able to determine what information to be studied to retrieve data from. This requires building rules and structure around the information to extract the critical elements.

Picking the data points that need to be analyzed: This helps to identify the data that needs to be analyzed. This depends upon the complexity of the data and the information, the extraction of that information and the calculation of the probability required can be straightforward or complex, but it is easy to determine by calculating the frequency, sometimes based upon the past analysis of similar data sources.

Identifying the key values from the extracted data set: Here one collates, identifies, and extracts the correct information from the larger gamut of data.

Extracting and Identifying Key Values: Learning techniques are more complex, and they rely on current and past data to produce a structure of past, valid experiences that can ultimately be compared to the new information and then interpreted and extracted. These steps help with both the extraction and

identification of the information that is extracted (points 3 and 4 from our step-by-step list).

TECHNIQUES OF DISCOVERING PATTERNS

There are various techniques of mining data. Different data mining techniques cater for different problems. The type of problem to be solved determines the type of data mining technique to employ for optimum results. Several authors classify data mining techniques in various ways. Some of these techniques are as enumerated (Adderley & Musgrove, 2001; Chen, Chung, Qin, Chau, Xu, Wang, Zheng & Atabakhsh, 2003; Brown, Pham & Vel, 2005)

Association Rule Learning: This method helps you to identify some interesting relations or dependency modelling between different variables in large databases. It helps to unpack some hidden patterns in the data that can be used to spot variables within the data and the concurrence of different variables that appear often in the dataset. Association rules are useful for examining and forecasting variables behaviour. Its usefulness is mostly felt or noticeable in the industry analysis. Association rule learning is used to determine variables clustering, and catalogue design (Pratte, 2001).

Anomaly or Outlier Detection: The observation for data items in a data set that does not match an expected pattern or behaviour is known as anomaly or outlier detection. Anomalies are known as outliers, novelties, noise, deviations or exceptions. They are source of critical and actionable information. An anomaly is an item that deviates considerably from the common average within a dataset or a combination of data. These types of items are statistically aloof when compared to the rest of the data and hence, indicates that something out of the ordinary has happened that requires additional attention. It is used in a variety of domains, such as intrusion detection, system health monitoring, fraud detection, fault detection, event detection in sensor networks, and detecting eco-system disturbances.

Visualization: This is the most useful technique used to discover data patterns. This technique is used at the beginning of the data mining process. It produces interesting projection of databases called Projection Pursuit. Visualization is a technique that converts poor data into good data letting different kinds of data mining methods to be used in discovering hidden patterns.

Neural Network: A neural network is a collection of interconnected neurons which could form a single layer or multiple layers. Neural networks are strong predictive modelling techniques.

Tracking patterns: One of the most basic techniques in data mining is learning to recognize patterns in your data sets. This is usually recognition of some aberration in your data happening at regular intervals, or an ebb and flow of a certain variable over time.

DATA MINING APPLICATIONS

The use of data mining is apparent today in developed countries and mostly by multinational companies and highly automated manufacturing companies. Companies with very high volume of transactions are embracing data mining and its allied technology. Various industries are now adopting data mining to their mission and critical business processes to gain competitive advantages and to help business grow (Rajkumar, 2014). Data mining applications are spreading into the areas of sales, marketing, banking, finance, healthcare, insurance, transportation, medicine, security issues and crime combating (Hand, Mannila & Smyth, 2001).

Data mining has been successfully utilized and its enormous benefits are enjoyed in various sectors. Some of this numerous benefits are as enumerated below (Karimabadi, Sipes, White, Marinucci, Dmitriev, Chao, Driscoll, & Balac, 2007).

In Sales and Marketing: Data mining enables businesses to understand the hidden patterns inside historical purchasing transaction data, thus helping in planning and launching new marketing campaigns in a prompt and cost-effective way.

In Banking and Finance: Data mining contributes to solving business problems by finding patterns, causalities, and correlations in business information and market prices that are not immediately apparent to managers because the volume of data is large.

In Insurance: Data mining is used in claims analysis such as identifying which medical procedures are claimed together; to forecasts which customers will potentially purchase new policies; to detect risky customers' behaviour patterns; and to detect fraudulent behaviour (Ogedengbe, Ogedengbe, Okhakhu & Ewanlen, 2018; Stolfo & Hershkop, 2005).

In Transportation: Data mining helps determine the distribution schedules among warehouses and outlets and analyze loading patterns. It shows where goods will be in short supply and when. It also shows the optimum route in the transportation scheduling and processes.

In Health Care and Medicine: Data mining facilitates the discovery of cures for diseases such as cancer and Parkinson's. It has already been used to discover how existing drugs can be used to treat other conditions.

In Agriculture: Data mining is used in predicting problems, disease detection and optimizing pesticides.

Crime and other vices: Crime is primarily the outcome of multiple adverse internal and external causes and conditions, such as biological, psychological, physiological, social, economic, educational, ethnical, environmental, seasonal, political, cultural and family conditions. Regardless of its complexity, to prevent crime it is vital to have an understanding of its roots. Data mining helps in crime prediction forecasting. Data mining is used in predicting future crime trends. This involves tracking crime rate changes from one year to another and to project those changes into the future. The basic method involves clustering of the individuals or states having the same crime trend and then using next year cluster information to classify the records (Yamuna & Bhuvaneswari, 2012).

DATA MINING IMPERATIVES

Data mining brings a lot of benefits to businesses, management science, society, governments as well as the individual. Data mining is not a simple process, and it relies on approaching the data in a systematic and mathematical fashion. But it also relies on being flexible, and taking data that might not necessarily fit into a nicely organized and sequential format. Data mining tools predict behaviours and future trends, allowing businesses to make proactive, knowledge driven decisions. Data mining discovers information that was not expected to be obtained.

In business, data mining helps to discover patterns and relationships in data that aids and helps in effective and efficient decision making. By using software to look for patterns in large batches of data, businesses can learn more about their customers to develop more effective marketing strategies, increase sales and decrease costs. Data mining helps in developing smarter marketing campaigns and to predict customer loyalty.

Succinctly put, in Businesses and Management data mining is used to discover patterns and relationships in order to help make better business and management decisions. Data mining can help spot sales trends, develop smarter marketing campaigns, and accurately predict customer loyalty. Data mining imperatives in business and management is summarize into and enumerated below.

Business Opportunities Creation: Data mining will allow businesses and managers spot potential business opportunities through pattern study and trend analysis, before they emerge.

Business/Investment Problem Eradication: Data mining has the ability to identify where we will likely run into business and investment problems so that we can avoid or totally eradicate them as quick as possible.

Business Relationship Management: Data mining has the capability to determine the patter and behaviour of data and this can be explored to manage relationships.

Market Segmentation: Data mining aids in identifying the common characteristics of customers who buy the same products from your company.

Customer: Data mining forecasts and predicts which customers are likely to dump a product for another or leave your company and go to your business competitor and how soon this may happen.

Fraud detection: Data mining quickly identifies which transactions are most likely to be fraudulent and which is likely to be bad and doubtful.

Direct marketing: Data mining helps to identify which prospects should be included in a mailing list to obtain the highest response rate.

Interactive marketing: Data mining predicts what each individual accessing a profile, record web site or catalogue, is most likely interested in seeing.

Market basket: Data mining and analysis give understanding of what products or services are commonly purchased together.

Trend analysis: Data mining reveals the difference between a typical customer current month and previous month.

Data mining applied in the context of law enforcement and intelligence analysis holds the promise of alleviating crime and crime related problems. Using a wide range of techniques it is possible to discover useful information to assist in crime matching, not only of single crimes, but also of series of crimes. A common technique used here is clustering, which is a model to anticipate crime trends. The results from this can potentially be used to lessen and even prevent crime for the forth coming years. Crime data mining has a promising future for increasing the effectiveness and efficiency of criminal and intelligence analyses (Stolfo, Lee, Chan, Fan & Eskin, 2001; Yamuna & Bhuvaneswari, 2012).

DRAWBACKS IN DATA MINING

Despite all the advantages and benefits of data mining, it should be noted that there are some caveats. There are some drawbacks, disadvantages and other issues of challenge in data mining. Some of these are carefully stated here.

The excessive work involved in data mining requires investment in new technology, high performance teams as well as staff training and development. There is difficulty in the data collection used in data mining and depending on the type of data needed; this can be an enormous work to be engaged in. Sometimes the necessary technology to carry out the data collection is not an easy task and could

consume many resources and that means high cost. Depending on the amount of databases, it may take some time to pre-process all the information needed from the mining of data.

Another drawback of data mining is that, the lack of an appropriate security system puts at risk the private information of the users. Data mining is not a perfect process, if the information is inaccurate, it would affect the outcome of the decision making process of the user concerned.

CONCLUSION

Prediction has always been the holy grail of all business endeavours. We have long had the ability to discern patterns in data in order to analyse in hindsight, and then apply this to future action. But data mining takes us a step closer to deciphering the unknown. Data mining gathers data from sundry sources and applying varying techniques transforms the data into useable information for business and management decision making. The two high-level primary goals of data mining, in practice, are prediction and description. While prediction involves using some variables or fields in the database to predict unknown or future values of other variables of interest, description focuses on finding human-interpretable patterns describing the data.

Today, data mining is apparent and its importance, relevance and imperatives are becoming more visible in all the spheres of human endeavours. Its applications are very glaring in developed countries and mostly by multinational companies and highly automated manufacturing companies. Companies with very high volume of transactions are embracing data mining and its allied technology. Various industries are now adopting data mining to their mission and critical business processes to gain competitive advantages and to help business grow. This study has highlighted how Data mining brings a lot of benefits to businesses, management science society, governments as well as the Individual.

In the same vein, some drawbacks of data mining were equally enumerated to help business owners, decision makers and managers guide and guard against the caveats in data mining, its processes and applications.

REFERENCES

Adderley, R. & Musgrove, P. B. (2001). Data mining case study: Modeling the behaviour of offenders who commit serious sexual assaults. Proceedings of the seventh ACM

SIGKDD International Conference on Knowledge Discovery and Data Mining; 215- 220

Brown, M. (2014). Steps to start data mining. http://scitechconnect.elsevier.com/5-steps-start-data- mining/ retrieved 19/12/2019.

Brown, B., Pham, B. & Vel, O. (2005). Design of a digital forensics image mining system. IIHMSP05, Melbourne

Chen, H., Chung, W., Qin, Y., Chau, M., Xu, J. J., Wang, G., Zheng, R. & Atabakhsh, H. (2003). Crime data mining: An overview and case studies. *ACM International Conference Proceeding Series*; *130*, 1–5.

Hand, D., Mannila, H. & Smyth, P., (2001). Principles of data mining. Cambridge, MA: MIT Press

Karimabadi, H., Sipes, T. B., White, H., Marinucci, M., Dmitriev, A., Chao, J.K., Driscoll, J. & Balac, N. (2007). *Journal of Geophysical Research, 112*, AA11215, doi:10.1029/2006JA012136.

Ogedengbe, F.A., Esther Ogedengbe, E. & Ovuworie, G.C. (2018). Economic implications of statistical quality control: Lessons for manufacturers in Nigeria. *American Journal of Operations Management and Information Systems 3*(2), 61-65

Ogedengbe, F. A., Ogedengbe, E., Okhakhu, C.O. & Ewanlen, D. O. (2018). Business forecasting imperatives in a depressed economy: A Nigeria case. *American Journal of Operations Management and Information Systems; 3*(2), 46-51

Rajkumar, P. (2014). Top 14 useful applications for data mining https://bigdata-madesimple.com/14-useful-applications-of-data-mining/.Retrieved 06/05/2019

Sainz, J. (2019). Data mining: The imperative reason for its existence.

https://medium.com/@jsainz3/data-mining-the-imperative-reason-for-its-existence-861000d7af9b. retrieved 19/12/2019.

Saltos, G & Cocea, M. (2017). An exploration of crime prediction using data mining on open data. *International Journal of Information Technology & Decision Making 16,*(5)

Stolfo, S. J. & Hershkop, S. (2005). Email mining toolkit supporting law enforcement forensic analyses. *ACM International Conference Proceeding Series, 89*, 221–222.

Stolfo, S. J., Lee, W., Chan, P. K., Fan, W. & Eskin. E. (2001). Data mining-based intrusion detectors: An overview of the Columbia IDS project. *ACM SIGMOD Record, 30*, 5–14. 4.

Vel, O., Anderson, A., Coney, M. & Mohay. G. (2001). Mining e-mail content for author identification forensics. *ACM SIGMOD Record. 30*(4).

Yamuna, S. & Bhuvaneswari, N.S. (2012). Data mining techniques to analyze and predict crimes. *The International Journal of Engineering and Science. 1*(2), 243-247