



# Role of Data Mining in Managerial Decisions

Dr. Javed Iqbal<sup>1</sup>

<sup>1</sup>Assistant Professor

School of Management Studies, BGSB University Rajouri, Jammu & Kashmir, India

javedkasana@gmail.com<sup>1</sup>

## ABSTRACT

In the age of technology, huge electronic data repositories are being maintained by business houses and financial institutions. Information for longer periods is being kept in these data repositories. The huge size of these data sources make it possible for financial analysts to come up with interesting information that helps in the decision making for future operations. In this paper an attempt has been made to analyse the role of data mining and its contribution to solving business problems in banking and finance by finding patterns, causalities, and correlations in business information and market prices that are not immediately apparent to managers because of the global market competition and market volatility. In this paper, the researcher has tried to highlight the application of data mining that has definitely positive impact on risk management, portfolio management, trading, customer profiling and customer care, where data mining techniques can be used in banks and other financial institutions to enhance performance through efficient decision-making.

**Keywords:** Data Mining, Financial Markets, Financial Institutions, Risk Management, Portfolio Management, Technology, Decision Making

## I. INTRODUCTION

The trio concepts of liberalisation, globalisation, and privatisation (LPG) is gaining an overall acceptance throughout the world economies, which is providing a framework for understanding and operating the minute issues confronting every individual, organisation or other institutions. These stakeholders somehow somewhere for managing their day to day operations come across the number of alternatives available to them in the environment. In other words, we can say that there is large number of information available in the external or internal environment which is relevant or possess some solutions for the problems that the organisations face. The problems or issues may be regarding their daily

generalisations or for systematic issues. In the present day world we have pertinent information available regarding concepts – like information regarding behaviour of customers in relation to the market change in terms of prices of goods and services, changing preferences, information about the intrinsic prices of shares and stocks, information regarding transportation, warehousing and logistics, and information regarding the changing trends in the workforce diversity, employees grievances, performance appraisal and so on and so forth. Thus, we can say that large portion of data is available but in the form of raw data. What usually is done is that, data is collected and refined in order to generate the essential information that suits the problems or issues that individuals or organisations face. Thus,

data available for problem solving cannot be accurate, since it needs refining. This data refining is known as "Data Mining". In simple words, data mining is defined as a process used to put out operational data from a larger set of any raw data. It implies analysing data patterns in large groups of data using one or more software. Data mining has uses in many fields, like science and research. As an application of data mining, businesses can learn more about their customers and develop more effective strategies related to various business functions and in turn leverage resources in a more optimal and insightful manner. This helps businesses to achieve long term and short term objective and make better decisions. Data mining involves effective data collection and warehousing as well as computer processing. For segmenting the data and evaluating the probability of future events, data mining uses sophisticated mathematical algorithms. Data mining is also known as Knowledge Discovery in Data (KDD).

## II. USES OF DATA MINING

The most pivotal example of Data Mining and Business Intelligence comes from service providers in the cell phone and utilities industries; they gather billing information, customer services interactions, website visits and other metrics to give each customer a probability score, then target offers and incentives to customers whom they perceive to be at a higher risk of churning. Next to mobile industry comes the retail world, Retailers target customers into three groups: 'Regency, Frequency, Monetary' (RFM) groups and target marketing and promotions to those different groups. A customer who spends little but often and last did so recently will be handled differently to a customer who spent big but only once, and also some time ago. The former may receive more loyalty, upsell and cross-sell offers, whereas the latter may be offered a fewer deals. One of the most recognized examples in field of data mining is in the world of e-commerce, Many E-commerce companies use Data Mining and Business Intelligence to offer cross-sells and up-sells through

their websites. One of the most famous of these is, of course, Amazon, who use sophisticated mining techniques to drive there People who viewed that product, also liked this' originality. There is a new emerging field, called Educational Data Mining, concerns with developing methods that discover knowledge from data originating from educational Environments. The goals of EDM are identified as predicting students' future learning behaviour, studying the effects of educational support, and advancing scientific knowledge about learning. Data mining can be used by an institution to take accurate decisions and also to predict the results of the student. With the results the institution can focus on what to teach and how to teach. Learning pattern of the students can be captured and used to develop techniques to teach them. Huge dollars have been lost to the action of fraudulent and pranksters. Old methods of fraud detection are time consuming and exhausting.

Data mining helps in providing meaningful patterns and turning data into information. Any information that is authentic and useful is knowledge. A perfect fraud detection system should protect information of all the users. A supervised method includes collection of sample records. These records are classified fraudulent or non-fraudulent. A model is built using this data and the algorithm is made to identify whether the record is fraudulent or not, which saves time and improves efficiency. With net banking everywhere, huge amount of data is generated with new transactions. Data mining can contribute in tackling business problems in banking and finance by finding patterns, causalities, and correlations in business information and market prices that are not immediately apparent to managers because the volume data is humungous or is generated too quickly to screen by experts. The managers may find this information for better segmenting, targeting, acquiring, retaining and maintaining a profitable customer base. History has proven that we have witnessed revolutionary changes in research. Data mining is helpful in data

cleaning, data pre-processing and integration of databases. Identification of any co-occurring sequences and the correlation between any activities can be known. Data visualization and visual data mining provide us with a clear view of the data. Apart from these uses, Patel highlighted the ten most common ways to use data mining:

1. Basket analysis
2. Sales forecasting
3. Database marketing
4. Merchandise planning
5. Card marketing
6. Call detail record analysis
7. Customer loyalty
8. Market segmentation
9. Product production
10. Warranties

### III. USERS OF DATA MINING

Following are some of the users of the data mining:

#### A. Banking and insurance

With analytic expertise, insurance companies can solve complex problems concerning fraud, compliance, risk management and customer attrition. Companies have used data mining techniques to price products more effectively across business lines and find new ways to offer competitive products to their existing customer base.

#### B. Networking

In a humongous market, where competition is cut throat, the answers are often within your consumer data. Multimedia and telecommunications companies use analytic models to make sense of mountains of customer's data, helping them predicting customer behaviour and offer highly targeted and relevant campaigns.

#### C. Knowledge Management

With unified, data-driven views of student progress, educators can predict student performance before

they set foot in the classroom – and develop intervention strategies to keep them on course. Data mining helps educators access student data, predict achievement levels and pinpoint students or groups of students in need of extra attention.

#### D. Secondary Sector (Manufacturing)

Aligning supply plans with demand forecasts is essential, as is early detection of problems, quality assurance and investment in brand equity. Manufacturers can predict wear of production assets and anticipate maintenance, which can maximize uptime and keep the production line on schedule.

#### E. Retailing

Large customer databases hold hidden insights that can help you improve customer relationships, optimize marketing campaigns and forecast sales. Through more accurate data models, retail companies can offer more targeted campaigns – and find the offer that makes the biggest impact on the customer.

### IV. CONCLUSION

Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions. Authors are strongly encouraged not to call out multiple figures or tables in the conclusion these should be referenced in the body of the paper.

### V. DATA MINING TECHNIQUES

According to Thearling (2002) the most widely used techniques in data mining are:

- Decision Trees: A decision tree is a decision tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. This is described as tree-shaped structures that rules for the classification of a data set. Examples of a decision tree methods are Chi-

square Automatic Interaction Detection (CHAID) and Classification and Regression Trees (CART).

- Rule Induction: The extraction of useful if-then rules from data based on statistical significance.
- Genetic Algorithms: He explained this to mean optimization techniques that use processes such as genetic combination, mutation, and natural selection in a design based on the concepts of evolution.
- Artificial Neural Networks: Non-linear predictive models that learn through training and resemble biological neural networks in structure.

## VI. PROBLEMS OR CHALLENGES OF DATA MINING

Data mining systems face a lot of problems and pitfalls. Data mining applications makes use of database to supply the raw data for input. This can also pose as a challenge and the challenges according to (Kamelsh Mhashilkar, 2011) are:

### A. Uncertainty

This refers to the error and degree of noise in the data. Data has got to be precise for consideration in a discovery system. Databases usually has the challenge of noise. Noise is defined as errors in either the values or attributes or class. According to (Kamelsh 2011: Ibid) missing data can be treated by discovery systems in a number of ways such as:

- Simply disregard missing values.
- Infer missing values from known values
- Omit the corresponding records.

### B. Distributed Data

Real world data is usually stored on different platforms in distributed computing environments. It could be in databases, individual systems, or even on the Internet. It is practically very difficult to bring all the data to a centralized data repository mainly due to organizational and technical reasons. For example, different regional offices might be having their own servers to store their data whereas it will not be feasible to store all the data (millions of terabytes)

from all the offices in a central server. So, data mining demands the development of tools and algorithms that enable mining of distributed data.

### C. Complex Data

The data available is actually heterogeneous and it could be multimedia data including images, audio and video, complex data, temporal data, spatial data, time series, natural language text and so on. It is really difficult to handle these different kinds of data and extract required information. Most of the times, new tools and methodologies would have to be developed to extract relevant information.

### D. Huge Database

Sometimes the information in the database is large enough that it becomes impossible to extract the relevant information because the data is large in quantity with mixed information of various disciplines.

## VII. ROLE OF DATA MINING IN MANAGERIAL DECISION MAKING

Data mining is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques. "Data mining sometimes called data or knowledge discovery. Data are any facts, numbers, or text that can be processed by a computer. Today, organizations are accumulating vast and growing amounts of data in different formats and different databases. Data mining is the process of analysing data from different perspectives and summarizing it into useful information the patterns, associations, or relationships among all this data can provide information.

## VIII. IMPACT OF DATA MINING IN MANAGERIAL DECISION MAKING

The concept of Data Mining is becoming increasingly popular as a business information management tool

where it is expected to disclose knowledge structures that can guide decisions in conditions of limited certainty. A data mining supports 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 mining allows the manager to perform more substantive, accurate and consistent analysis.

The goal of a data mining 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. To be successful, 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. 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.

There are different areas where managers use data mining in order to extract the information relevant to their areas of operations, for example:

1. Information regarding financial accounting
2. Areas of marketing management
3. Recruitment of human resources
4. Information in the areas of production management
5. Information about the behaviour of employees in an organisation
6. Logistics management
7. Data in the areas of research and development

## VII. CONCLUSION

It is concluded that one of the key concerns during this analysis is that of the data's quality. High-quality data is important because it gives us accurate and timely information to manage services and accountability. Data mining helps the businesses to handle the situations of market change, volatility in the prices, technological changes etc. Also, it helps managers to prioritize and ensure the best use of resources. This is a no brainer that good quality data will lead to valuable information and appropriate insights for your organization. But, obtaining high-quality data is not an easy task. Improving your data quality and sustaining good quality data output is one of the major challenges faced by the enterprises today.

According to IBM, data-quality related problems can result in a loss of millions of dollars in revenue. Wrong decisions are made due to poor quality of data and poor data management processes. Due to this, many companies lose their customers and clients. Thus, if data quality is not ensured, your data can become a risky liability instead of a significant asset.

## IX. SUGGESTIONS

Some of the suggestions will help the managers to improve the quality of their information in order to make the effective decisions that will enable them to achieve the long term as well as short term goals.

1. Individual, business organisations, and other Companies should identify the right and reliable data sources, both internally and externally, to improve the quality of incoming data. Incorrect data may result from migration of data from one database to another.

2. A manager should keep an eye on data overloading and must ensure that the data entry team is not under pressure to deliver expected results. If data entry specialists are overloaded with work, they can become tired and that can result in data entry errors, which can remain consistent through the time.

3. A manager should adopt statistical analysis of the raw data by applying the known statistical tools like Mean, Median, Mode, and chi-square test, Z-test, T-test and F-test etc.

4. Companies should review the data as reviewing is an efficient way to check the correctness of the data. Companies must incorporate an efficient way to review and double check the data entered. It is always beneficial to hire a team of quality assurance professionals who can review the data and help in reducing the data errors to a large extent.

5. Having a good and healthy work environment helps the employees make lesser mistakes and therefore has a direct impact on data accuracy. Companies need to provide a healthy work environment to their data entry professionals which help in retaining their focus.

6. Companies must adopt highly robust data entry quality standards such as matching, geo-coding, data monitoring, data profiling, linking, etc. This ensures

that the data entered conforms to pre-defined data standards which in turn help in improving the data quality.

## X. REFERENCES

- [1] Wang, S. (2010). A Comprehensive Survey of Data Mining-Based Accounting-Fraud Detection Research. International Conference on Intelligent Computation Technology and Automation, vol. 1, pp.50-53, 2010.
- [2] Bologna, Jack & Lindquist, R. J. (1987). Fraud Auditing and Forensic Accounting. New York: John Wiley & Sons, 1987.
- [3] Feroz, E.H., Kwon, T.M., Pastena, V., & Park, K.J. (2000). The efficacy of red flags in predicting the SEC's targets: an artificial neural networks approach, International Journal of Intelligent Systems in Accounting, Finance, and Management 9 (3) 145–157.
- [4] T.G. Calderon and J.J. Cheh: "A Roadmap for Future Neural Networks Research in Auditing and Risk Assessment", International Journal of Accounting Information Systems, Volume 3, Issue 4, December, 2002, pp.203-236.
- [5] M.J. Kim and I. Han: "The Discovery of Experts' Decision Rules from Qualitative Bankruptcy Data using Genetic Algorithms", Expert Systems with Applications, Volume 15, Issue 4, November, 2003, pp.637-646.
- [6] C.S. Park and I. Han: "A Case-based reasoning with the Feature Weights Derived by Analytic Hierarchy Process for Bankruptcy Prediction", Expert Systems with Applications, Volume 23, Issue 3, October, 2002, pp.255-264.
- [7] M.J. Beynon and M.J. Peel: "Variable Precision Rough Set Theory and Data Discretisation: an Application to Corporate Failure Prediction", Omega International Journal of Management Science, Volume 29, Issue 6, December, 2001, pp. 561-576.