Data Mining: A basis to Decision Theory

Authors Name: Rupal Mathur [1],Mridul Kumar Mathur[2], Pallavi Upadhyay [3]

[1]Asst. Prof. Department of Management Studies, Lachoo Memorial College of Science and Technology, Jodhpur, Rajasthan, India
[2] Asst. Prof. Department of Computer Science, Lachoo Memorial College of Science and Technology, Jodhpur, Rajasthan, India
[3] Asst. Prof. Department of Computer Science, Lachoo Memorial College of Science and Technology, Jodhpur, Rajasthan, India
mathur_mridul@yahoo.com[1], mubit@yahoo.com[3]

Abstract
Decisions are basis to every business, and to take the decisions we require having both qualitative and quantitative information processing at large. Decision making requires a crisp set of processed information and as the business grows or horizon expands more and more requirement of information and previous decision basis are needed. To inculcate this, we need computer as a tool for processing of information. But for processing requirements data set should be well formatted. Data mining is one of the basic concepts of adhering to produce relevant information for decision making. In this paper the authors proposes a model of data mining for decision making.

Introduction
The word "Globalization" has come to dominate the world. Globalization has bought new opportunities for the businesses. Exploration of new markets, technology and customers has become the prime research area for the business houses. We moved from small shops to big malls. Businessmen are now calculating millions dollars of profits. With the development in technology and their introduction in the global markets, there is not only a sound increase in the demand for commodities but has also led to greater utilization. More and more people are now connected to the world's trade happenings through computers. Buyers of products and services in all nations comprise one huge group who gain from world trade for reasons encompassing opportunity charge, comparative benefit, economical to purchase than to produce, trade's guidelines, stable business and alterations in consumption and production.
With the expansion in firm's operations in different countries, gathering information of external as well as internal environment has become a crucial task for the managers. Along with information related to stakeholders (Suppliers, customers, investors, etc) also turning into large and complex databases.

Information
Good decisions are based on good information and experience in interpreting information.
With the expansion in firm's operations in different countries, gathering information of external as well as internal environment has become a crucial task for the managers. Along with that the information related to stakeholders (Suppliers, customers, investors, etc) is turning into large and complex databases.
Information plays a crucial role in decision making as it is the basis for any decision. Historic data, research results, intuitions, predictions, market analysis, etc. can become a key source when decision is taken.

Decision Theory

Almost every business activity involves decisions. Therefore, to theorize about decisions is almost the same as to theorize about business activities. A decision making is required when numbers of alternatives are available to choose between. With every alternative certain events are associated, influenced by certain parameters which are uncontrollable in nature having their impact on each alternative with different ratio.

In the situations treated by decision theorists, there are options to choose between, and we choose in a non-random way. Our choices, in the situations, are goal-directed activities. Hence, decision theory is concerned with goal-directed behavior in the presence of options.

The decision process is divided into the following five steps:

1. Identification of the problem.
2. Obtaining necessary information.
3. Production of possible solutions.
4. Evaluation of such solutions.
5. Selection of a strategy for performance.
6. Implementation of the decision.

In decision-making, preference relations are used to find the best alternative. The following simple rules can be used for this purpose:

1. An alternative is best if and only if it is better than all other alternatives. If there is a uniquely best alternative, choose it.
   There are cases in which no alternative is uniquely best, since the highest position is "shared" by two or more alternatives.
2. An alternative is (among the) best if and only if it is at least as good as all other alternatives. If there are alternatives that are best, pick one of them.
   Decision theory provides the best suitable technique with the combination of alternatives/options and events (uncontrollable parameters) and tools for comparison between all possible alternatives in both cases i.e. under perfect information and imperfect information (uncertainty and risk).

What is data mining good for?

Data mining software allows users to analyze large databases to solve business decision problems. Data mining is, in some ways, an extension of statistics, with a few artificial intelligence and machine learning twists thrown in.

Data mining is not a business solution, it is just a technology. Data mining, extracts information from a database that the user did not know existed. Data mining is computer-aided pattern which discovers previously unknown inter-relationships and recurrences across seemingly unrelated attributes in order to predict actions, behaviors and outcomes. Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. These tools can include statistical models, mathematical algorithms, and machine learning methods.

Data Mining and Decision Making

Decisions are made based upon the availability of data as well as the decision situation that is environment. Basically there are three types of decision making environment: certainty, uncertainty and risk.

Decision-making under certainty: In this case the decision maker has the complete knowledge of consequence of every decision choice with certainty. Obviously he/she will select an alternative that yields the largest return (payoff) for the known future. In this decision

model, certainty means that only one possible state of nature (future) exists.

Decision-making under risk: In this case the decision-maker has less than the complete knowledge with certainty of the consequences of every decision choice. This means there is more than one state of nature and for which he /she makes an assumption of the probability with which each state of nature will occur.

Decision making under uncertainty: In this case the decision maker is unable to specify the probabilities with which the various states of nature will occur. Thus, decisions under uncertainty are taken with even less information than decisions under risk. Most of the business decisions are based on the future planning of business operations, depending on the timeframe and seasonality factors. Data mining tools help in formulation of both strategic and tactical decisions that are necessary for an organization to survive in the business, therefore data mining tools can be regarded as the best solution to decision making under risk and uncertainty.

The authors describes the implementation of data mining in the decision making process to boost the quality of the decisions linked with uncontrollable parameters. It has been suggested that the new information must be compared with the past records matching the same activity. This results in refining massive quantities of data to deliver relevant and actionable business information.

The second step i.e. information gathering of decision making process is the most crucial phase. If irrelevant data or the incomplete knowledge of uncontrollable parameters becomes the basis of decision making process then the whole goal oriented decision making fails and also the firm has to bear the loss of the cost incurred for collecting such information.

The decisions are highly influenced by:

1. The information gathered from historic data, researches, etc which act as a guideline for taking new decisions.
2. The uncontrollable parameters, the deviation of which can change the preference of the decision maker while choosing the best alternative.

The formal structure of a decision problem in any area can be put in four parts:

1. The choice of an objective function defining the relative interest of different outcomes.
2. Specification of the policy alternatives which are available to the decision maker.
3. specification of the model i.e. empirical relations that link the objective function, or the variables that enter into it, with the policy alternatives and possibly other variables
4. Computational methods for choosing among the policy alternatives that one which performs best as measured by the objective function.

Data mining is the best computational method which measures the relevancy of the information with the objective function. ***The extraction of hidden predictive information from large databases*** is a powerful new technology with great potential to help the companies to describe the information market of tomorrow.

A model should be developed with the combination of data mining and decision theory so that a relationship can be established between information collected through market survey or research and the data of past decisions made, retrieved by data mining.

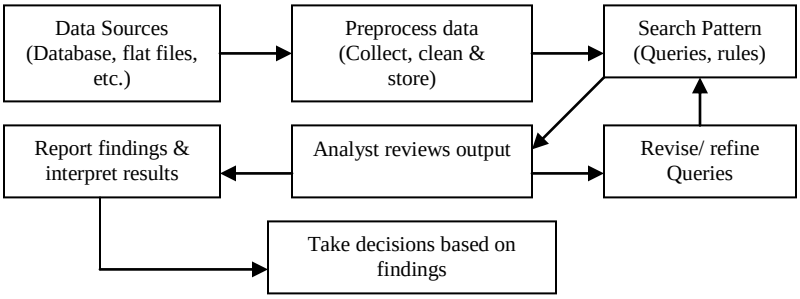

Figure 1: Basic model of Data Mining Process

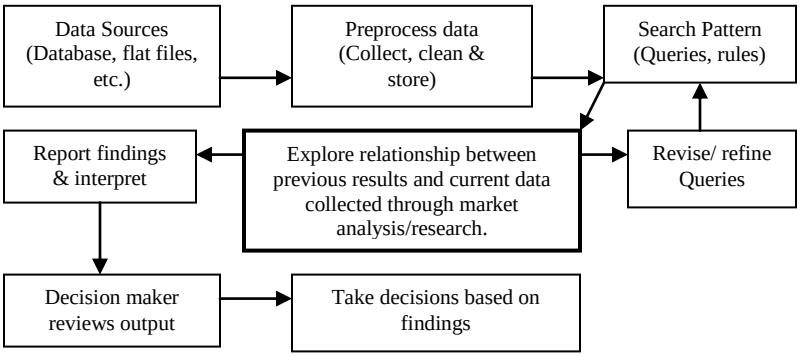

Figure 2: Proposed model of Data Mining Process

Conclusion

Decision making is the vital job for the managers as the success of any organization depends on the quality of the decisions made.

Data mining tools predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions. Data mining is concerned with learning from data and transformation of that data into useful information, data mining helps to find out the patterns and associations between the variables in the data. If the relationship is established between the past record and the data from the new research then the base for decision making becomes strong and also the uncontrollable elements can be analyzed effectively.

All national and multinational companies are using various data mining techniques and analytics to explore the relationship between various micro and macro economic factors and make policies using data analytics.

References:

[1] Data Intelligence Group, Pilot Software, "An Overview of Data Mining at Dun & Bradstreet", DIG White Paper 95/01, September 1995.

[2] I .Krishna Murthy, "Data Mining- Statistics Applications: A Key to Managerial Decision Making".

[3] Jiawei Han, Micheline Kamber, "Data mining: concepts and techniques", second edition.

[4] Taha A Hamdy, "operations Research: An Introduction", Pearson education, 8/E, 2008.

[5] Sharma J.K, "Operations Research: Theory and Applications ",Macmillan, 4/E, 2009.