

A Knowledge discovery Approach in Shopping Complex Database (ASCD)

K.Kavitha
Assistant Professor
Ramanathan College for Women
Karaikudi.

Dr. E. Ramaraj
Director , Computer Centre
Alagappa University,
Karaikudi.

Abstract:

Data mining and Knowledge Discovery (KD) has been widely accepted as a key technology for enterprises to improve their abilities in data analysis, decision support and the automatic extraction of knowledge from data. Existing method has framed the process of information extraction and also referred to as the knowledge discovery process as a series of strategic search decisions, subject to constraints with the objective of attaining a sufficient level of domain specific knowledge for use in enterprise planning. Prediction in financial domains is absolutely difficult for a number of reasons. Existing theories tend to be weak or non-existent, which makes problem formulation open, ended by forcing us to consider a large number of independent variables and thereby increasing the dimensionality of the search space. In this paper we are providing an effective association rule mining for developing the super market industries. We hope this would be the great help to them.

Keywords:

Data mining, Knowledge Discovery, Decision Making.

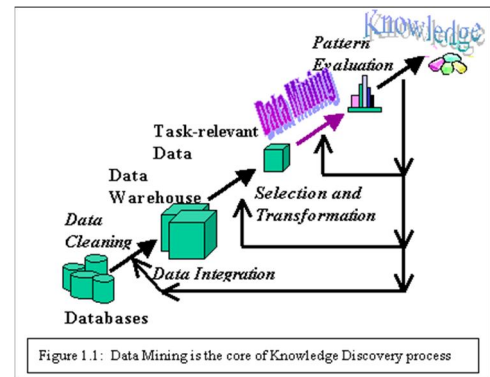
Introduction:

Across a wide variety of fields, data are being collected and accumulated at initial pace. There is an emergency need for a new generation of computational theories and tools to assist peoples in extracting useful information from the rapidly growing volumes of digital database. These theories and tools are the subject of the emerging field of knowledge discovery in databases (KDD). At an abstract level, the KDD field is concerned with the development of methods and techniques for making raw data into useful information. The basic problem addressed by the KDD process is one of mapping low level into other forms that might be more compact more abstract or more useful at the core of the process is the application of specific data mining methods for pattern discovery and extraction. This paragraph begins by discussing the historical idea of KDD and data mining and their usage with other related fields.

A brief summary of recent KDD real world applications is provided. Definitions of KDD and data mining are provided and the general multistep KDD process is outlined. This multistep process has the application of data mining algorithms as one particular step

in the process. The data-mining step is discussed in more detail in the context of specific data-mining algorithms and their application. Real world practical application issues are also outlined. Our basic assumption in predicting financial markets is that it is not possible to do so most of the time. This is consistent with remarks many financial professionals have made.

In particular many trading professionals and vendors do feel that there are times admittedly few where they can predict relatively well. This philosophy, “generally agnostic but occasionally making bets,” has important implications for how we approach the modelling problem. One of the major challenges is to reduce the noisy periods by being more selective about the conditions under which to invest to find patterns that offer a reasonable number of opportunities to conduct high risk-adjusted-return trades. In doing so, we must consider explicitly the trade-off between model coverage and model accuracy. Trying to give an accurate prediction for all data points is unlikely to succeed. On the other must be accurate enough and simultaneously general enough to allow sufficient high-probability opportunities to trade effectively.



The use of very huge databases has enhanced in the last years from supporting transactions to additionally reporting business problems and trends. The interest in analysing the data has increased due to business competition. One important topic is customer relationship management with the particular tasks of customer segmentation, customer profitability and customer retention and customer acquisition. Other tasks are the prediction of sales in order to minimize stocks, the prediction of electricity consumption or telecommunication services at particular day times in order to minimize the use of external services or optimize network routing, respectively.

The medicine sector demands several analysis tasks for resource management, quality control and decision making. Existing databases which were designed for transactions, such as billing and booking are now considered a mine of information and digging knowledge from the already gathered data is considered a tool for building up an organizational memory. Managers of an institution want to be informed about status and trends of their business. Hence, they

demand concise reports from the database department.

Online Analytical Processing offers interactive data analysis by aggregating data and counting the frequencies of occurrence.

This already answers questions like the following:

- What are the attributes of my most frequent customers?
- Which are the frequently sold products?
- How many returns did I receive after my last direct mailing action?
- What is the average duration of stay in my hospital?

Proposed Method:

Data mining methods and techniques can predict information that many traditional business analysis and statistical techniques fail to deliver. Additionally, the application of data mining techniques further yields the value of data warehouse by converting expensive volumes of data into valuable assets for future tactical and strategic business enhancement. Management information systems (MIS) should provide advanced capabilities that give the user the power to ask more sophisticated and pertinent questions. It empowers the right people by providing the specific information they need.

Most of the researcher doing in the following domains like

- Association Rule Discovery.

- Classification.
- Clustering.
- Sequential Pattern Discovery.
- Regression.
- Deviation Detection

But majority of researcher are doing in first three methods only. In our research also we have chosen association rule discovery method. We have tried to improve the shopping complex business strategy due to trends in fashion world. We have collected the database of purchasing details, billing details from the point of sale. Some of the stores using bar code reader system to recognize the products. In this case we have to collect the point of sale bar code details. In this we have to apply the association rule discovery method to find out the similarity of the purchase.

In our survey we have collected lot of information from the customer and vendors of shopping mall. They have required which of combination customers buying more? To find out the solution we have undertaken this automatic research application to improve the business strategy. In appendix 1 we have given the sample data set and association rule what we have driven from the dataset. From the driven rules we can predict some solution based on that vendors can arrange the products and get maximum profit for their business and also they can get business enhancement if they are not having that items. The same rule we can specify for specific products also like which brand customers more like? , which combination customers most like?

Conclusion:

In this paper we have provided an efficient method for developing the business strategy. This method will help to shopping mall and used to enhance the business strategy. This is the combination approach of data mining and real time business strategy. We hope this will give better result in the industries side.

Future work:

As initiation process we have undertaken the association rule mining process. In future we will try with the clustering and classification methods to improve the solution. Now we have tried just association rule mining later we will include the clustering and classification technique to improve the efficiency of result.

Reference:

[1] P. Bollmann-Sdorra, A.M. Hafez and V.V. Raghavan, "A Theoretical Framework for Association Mining based on the Boolean Retrieval Model," DaWaK 2001, September 2001.
 [2] A.M. Hafez, "Association mining of dependency between time series," *Proceedings of SPIE* Vol. 4384, SPIE AeroSense, April 2001.
 [3] V.V. Raghavan and A.M. Hafez, "Dynamic Data Mining," IEA/AIE 2000, pp.220-229, 2000.
 [4] A.K. Tung, J. Han, L.V. Lakshmanan and R.T. Ng, "Constraint-based Clustering in Large Databases," Proc. Int. Conf. on Database Theory, pp. 405-419, 2001.
 [5] E Turban, 1997. Decision support systems and expertsystems (Fifth Edition), Prentice-Hall, London: UK.
 [6] HL Viktor, 1999. *Learning by Cooperation: An Approach to Rule Induction and Knowledge Fusion*, PhD dissertation, Department of Computer Science, University of Stellenbosch, Stellenbosch: South Africa.
 [7] Abraham Bernstein, Shawndra Hill, and Foster Provost. An Intelligent Assistant for the Knowledge Discovery Process. Technical Report IS02-02, New York University, Leonard Stern School of Business, 2002.
 [8] Joerg-Uwe Kietz, Regina Zuecker, Anna Fiammengo, and Giuseppe Beccari. Data Sets, Meta-data and Preprocessing Operators at Swiss Life and CSELT. Deliverable D6.2, IST Project MiningMart, IST-11993, 2000.
 [9] Pavel Brazdil. Data Transformation and Model Selection by Experimentation and Meta-Learning. In C.Giraud-Carrier and M. Hilario, editors, *Workshop Notes – Upgrading Learning to the Meta-Level: Model Selection and Data Transformation*, number CSR-98-02 in Technical Report, pages 11–17. Technical University Chemnitz, April 1998.

Appendix 1

| Transaction ID | Items |
|----------------|----------------------------|
| MT12 | Shampoo, Conditioner, comb |
| MT13 | Shampoo, lotion |
| MT14 | Lotion, Conditioner, soap |
| MT15 | Soap, oil, lotion |
| MT16 | Lotion, oil, detergent |

Some of the Discovered Rules are:

| |
|-----------------------------------|
| Shampoo, conditioner -> -> Lotion |
| Conditioner, oil -> -> shampoo |
| Shampoo -> -> lotion |