# MINING A DATA IN E-COMMERCE

**Dr. A. M. Sote**
Assistant Professor,
Department of Computer Science,
Arts, Commerce and Science College,
Arvi (Wardha).
E-mail : amar_sote@rediffmail.com

-----------------------------------------------------------------------------------------------------------------

*Abstract: Electronic commerce processes and data mining tools have revolutionized many companies. Data that businesses collect about customers and their transactions are the greatest assets of that business. Data mining is a set of automated techniques used to extract hidden or previously unknown pieces of information from large databases, using different criteria, which makes it possible to discover patterns and relationships. This paper discusses the important role of business based on data mining knowledge development to detection the relation of data mining and electronic commerce.*

Keywords:  e-commerce, data mining, business intelligence.

-----------------------------------------------------------------------------------------------------------------

**Introduction :**

In today's business world there is an abundance of available data and a great need to make good use of it. In the first, data must be organized by data base tools and data warehouses, and then it needs an instrument for knowledge discovery. Data mining can be defined as the art of extracting non- obvious, useful information from large databases. This emerging field brings a set of powerful techniques which are relevance for companies to focus their efforts in taking advantage of their data.

Data mining tools generate new information for decision makers from very large databases. The various mechanisms of this generation include abstractions, aggregations, summarizations, and characterizations of data [1]. These forms, in turn, are the result of applying sophisticated modelling techniques from the diverse fields of statistics, artificial intelligence, database management and computer graphics.

Having a huge amount of data, make some problems for detection of hidden relationships among various attributes of data and between several snapshots of data over a period of time. These hidden patterns have enormous potential in predictions and personalization in e-commerce. Data mining has been pursued as a research topic by at least three communities: the statisticians, the artificial intelligence researchers, and the database engineers [2].

Although much work has been done to date, more studies need to be conducted to as various subjects in a variety of e-commerce problems. The purpose of this paper is a present of data mining methods and expression application of data mining in business.

## 2. APPLICA TIONS OF DATA MINING IN E-COMMERCE :

In this section, we had discussed salient applications of data mining techniques in e commerce for making a business more convenient.

### 2.1 Customer Profiling :

It may be observed that customers drive the revenues of any organization. Acquiring new customers, delighting and retaining existing customers, and predicting buyer behaviour will improve the availability of products and services and hence the profits. Thus the end goal of any data mining exercise in e-commerce is to improve processes that contribute to delivering value to the end customer. Consider an on-line store like http:www.dell.com where the customer can configure a PC of his/her choice, place an order for the same, track its movement, as well as pay for the product and services. With the technology behind such a web site, Dell has the opportunity to make the retail experience exceptional. At the most basic level, the information available in web log files can detect what prospective customers are seeking from a site.

Companies like Dell provide their customers access to details about all of the systems and configurations they have purchased so they can incorporate the information into their capacity planning and infrastructure integration. Back-end technology systems for the website include sophisticated data mining tools that take care of knowledge representation of customer profiles and predictive modelling of scenarios of customer interactions. For example, once a customer has purchased a certain number of servers, they are likely to need additional routers, switches, load balancers, backup devices etc. Rule-mining based systems could be used to propose such alternatives to the customers.

### 2.2 Recommendation Systems :

Systems have also been developed to keep the customers automatically informed of important events of interest to them. The article by Jeng & Drissi [3] discusses an intelligent framework called PENS that has the ability to not only notify customers of events, but also to predict events and event classes that are likely to be activated by customers. The event notification system in PENS has the following components: Event manager, event channel manager, registries, and proxy manager. The event-prediction system is based on association rule- mining and clustering algorithms. The PENS system is used to actively help an e-commerce service provider to forecast the demand of product categories better. Data mining has also been applied in detecting how customers may respond to promotional offers made by a credit card e-commerce company [4]. Techniques including fuzzy computing and interval computing are used to generate if-then-else rules. Niu et al present a method to build customer profiles in e-commerce settings, based on product hierarchy for more effective personalization [5]. They divide each customer profile into three parts: basic profile learned from customer demographic data; preference profile learned from behavioural data, and rule profile mainly referring to association rules.

Based on customer profiles, the authors generate two kinds of recommendations, which are interest recommendation and association recommendation. They also propose a special data structure called profile tree for effective searching and matching.

## 2.3 Web Personalization :

Mobasher presents a comprehensive overview of the personalization process based on web usage mining [6]. In this context, the author discusses a host of web usage mining activities required for this process, including the pre-processing and integration of data from multiple sources, and common pattern discovery techniques that are applied to the integrated usage data. The goal of this paper is to show how pattern discovery techniques such as clustering, association rule-mining, and sequential pattern discovery, performed on web usage data, can be leveraged effectively as an integrated part of a web personalization system. The author observes that the log data collected automatically by the Web and application servers represent the fine-grained navigational behaviour of visitors.

Depending on the goals of the analysis, e-commerce data need to be transformed and aggregated at different levels of abstraction. E-commerce data are also further classified as usage data, content data, structure data, and user data. Usage data contain details of user sessions and page views. The content data in a site are the collection of objects and relationships that are conveyed to the user. For the most part, the data comprise combinations

## 2.4 Customer Behaviour in E-commerce :

For a successful e-commerce site, reducing user-perceived latency is the second most important quality after good site-navigation quality. The most successful approach towards reducing user-perceived latency has been the extraction of path traversal patterns from past users access history to predict future user traversal behaviour and to prefect the required resources. However, this approach is suited for only non-e-commerce sites where there is no purchase behaviour. Vallamkondu & Gruenwald describe an approach to predict user behaviour in e-commerce sites [7]. The core of their approach involves extracting knowledge from integrated data of purchase and path traversal patterns of past users (obtainable from web server logs) to predict the purchase and traversal behaviour of future users.

Web sites are often used to establish a company's image, to promote and sell goods and to provide customer support. The success of a web site affects and reflects directly the success of the company in the electronic market. Spiliopoulou & Pohle propose a methodology to improve the success of web sites, based on the exploitation of navigation-pattern discovery [8]. In particular, the authors present a theory, in which success is modelled on the basis of the navigation behaviour of the site's users. They then exploit web usage mining (WUM), a navigation pattern discovery miner, to study how the success of a site is reflected in the users' behavior. With WUM the authors measure the success of a site's components and obtain concrete indications of how the site should be improved. The data sources used to deliver or generate data include static HTML/XML pages, images, video clips, sound files, dynamically generated page segments from scripts or other applications, and collections of records from the operational database(s).

Site content data also include semantic or structural metadata embedded within the site or individual pages, such as descriptive keywords, document attributes, semantic tags, or HTTP variables. Structure data represent the designer's view of the content organization within the site. This organization is captured via the inter-page linkage structure among pages, as reflected through hyperlinks. Structure data also include the intra-page structure of the content represented in the arrangement of HTML or XML tags within a page. Structure data for a site are normally captured by an automatically generated site map which represents the hyperlink structure of the site. The operational database(s) for the site may include additional user profile information. Such data may include demographic or other identifying information on registered users, user ratings on various objects such as pages, products, or movies, past purchase or visit histories of users, as well as other explicit or implicit representations of users' interests.

**Conclusions :**

In this paper, we have presented how web mining (in a broad sense, data mining applied to ecommerce) is applicable to improve the services provided by e-commerce based enterprises. Specifically, we first discussed some popular tools and techniques used in data mining. Statistics, AI and database methods were surveyed and their relevance to data mining in general was discussed. We then presented a host of applications of these tools to data mining in e-commerce. Later, we also highlighted architectural and implementation issues. We now present some ways in which web mining can be extended for further research. With the growing interest in the notion of semantic web, an increasing number of sites use structured semantics and domain ontology as part of the site design, creation, and content delivery. The notion of Semantic Web Mining was introduced by Berendt et al [9]. The primary challenge for the next-generation of personalization systems is to effectively integrate semantic knowledge from domain ontology into the various parts of the process, including the data preparation, pattern discovery, and recommendation phases. Such a process must involve some or all of the following tasks and activities [6].

**References :**

- P. L. Carbone, "Expanding the meaning of and applications for data mining," Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, vol. 3, pp. 1872-1873,2000.

- N. R. S. Raghavan, "Data mining in e-commerce: A survey," Sadhana – Academy Proceedings in Engineering Sciences, vol. 30, pp. 275-289, 2005.

- J.-1. Jeng and Y. Drissi, "PENS: A predictive event notification system for e-Commerce environment," Proceedings - IEEE Computer Society's International Computer Software and Applications Conference, pp. 93-98, 2000.

- X. Z. Zhang, "Building personalized recommendation system in E-Commerce using association rule-based mining and classification," in Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, ICMLC 2007,pp.4113-4118,2007.

- L. Niu, X. W. Van, C. Q. Zhang, and S. C. Zhang, "Product hierarchy-based customer profiles for electronic commerce recommendation," Proceedings of 2002 International Conference on Machine Learning and Cybernetics, vol. 2, pp. 1075- 1080,2002.

- B. Mobasher, "Web usage mining and personalization," Practical Handbook of Internet Computing, 2004.

- S. Vallamkondu and L. Gruenwald, "Integrating purchase patterns and traversal patterns to predict http requests in e-commerce sites," IEEE Int. Conj on e- commerce, pp. 256-263, 2003.

- M. Spiliopoulou and C. Pohle, "Data mining to measure and improve the success of web sites," J. Data Mining and Knowledge Discovery, 2000.

- B. Berendt, A. Hotho, and G. Stumme, "Towards semantic web mining," Proceedings of the International Semantic Web Conference, vol. 2342, pp. 264-278, 2002.