

# Knowledge Management using Data Mining in E-Commerce

Pallavi Sood

*PG Department of Computer Science & IT, Lyallpur Khalsa College, Jalandhar, Punjab, India*

**Abstract—** In today's competitive world, people are flooded with huge amount of data which has filled almost every aspect of our lives. Our abilities of both generating and gathering data have been expanding speedily. This rapid growth of available data volume is a result of the computerization of our society and the fast development of powerful data collection and storage tools. We need to extract useful information from this large amount of data for better decision making. The solution to analyze huge amount of data and extract useful information from it is Data Mining popularly referred to as knowledge discovery from data (KDD). Knowledge management provides systematic approach to captures, structures, maintains, disseminate and access to the knowledge within information systems. This paper describes how data mining techniques helps for knowledge management in e-commerce.

**Index Terms:** data mining, knowledge management, knowledge discovery, e-commerce.

## I.INTRODUCTION

Large volume of structured, semi-structured and unstructured data, which is called big data, is available nowadays. This data provides big advantages for those organizations which use electronic commerce. Electronic commerce (e-commerce) refers to the buying or selling of goods or services via electronic channels over the internet. Electronic channels include computers, tablets, smart phones and other smart devices.[1] The data is gathered from buyers, sellers, business and market environment. From this gathered data, extracting the useful knowledge which helps in smooth running of business is very important. This extraction of knowledge is provided by data mining. Many e-commerce sites like Amazon, Flipkart, firstcry use data mining for cross-selling and upselling or promotion for their products to show "Buy it with, "You may also like", "Frequently bought together".

These recommendations are provided using knowledge gathered from data mining over the purchasing or visiting history of the customers of the website.

In the paper, I mention that the data mining plays an important contribution to knowledge management. The purpose is to show how data mining techniques can be used for managing knowledge in ecommerce, which would lead to a better and smooth running of business. Data mining is believed to be a good promoter of e-commerce. The challenge of extracting knowledge from data draws by integrating statistics, databases, pattern recognition and artificial intelligence together for the purpose of better decision making.

The paper is organized as follows: the next section discusses about data mining and process of data mining. The following section presents a review of knowledge management and types of knowledge. The next section discusses about technologies used for knowledge discovery in data mining. The final part provides a conclusion and discussion about future work.

## II.DATA MINING

We are deluged by variety of data like scientific data, medical data, demographic data, financial data, marketing data, social media data etc. People only want useful information that they need for their decision making process. So analyzing large amount of data is a necessity. Every business advantages from collecting and analyzing its data. Data mining turns a huge amount of data into knowledge. Data mining techniques helps us for discovering interesting patterns from data in various applications. Data mining refers to extracting knowledge from huge volume of data.[2] Data mining can be suitable to any kind of data repository or storage as well as transient data such as data streams. Data repositories

include relational databases, data warehouses, advanced database systems such as object relational databases and specific application oriented databases like spatial databases, time series databases, text databases and multimedia databases. Data mining techniques are adopted by large scale ecommerce businesses to aid with marketing and product development. In ecommerce data mining there are three important processes that are followed to turn data into knowledge.

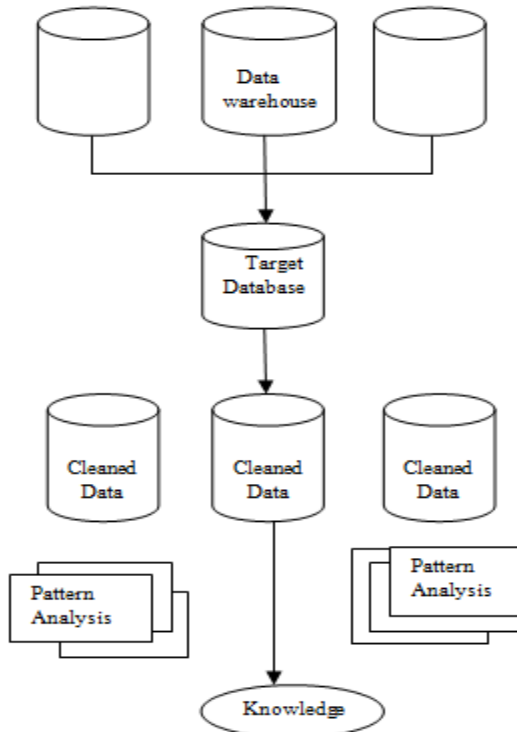


Fig. 1 Data mining process in e-commerce

- a. Data Selection: At this stage, data that are relevant to the analysis task are retrieved from the collected data.
- b. Data Transformation: In this phase, selected data are transformed into forms which are appropriate for data mining procedure by performing summary or aggregation operations. Here one type of data is converted to another according to need. At this stage, strategies also defining to handle missing data.
- c. Data Mining: It is an essential process where intelligent techniques are applied to extract data patterns of interest.
- d. Result interpretation and validation: For better understanding of data, interesting patterns that

represents knowledge are identified based on given measures. For validity, the robustness is checked by data mining application test.

- e. Incorporation of the discovered knowledge: It is the final step where the result of discovered knowledge is represented to decision makers so that new discovered patterns can be applied for better decision making.

### III.KNOWLEDGE MANAGEMENT AND KNOWLEDGE MANAGEMENT TECHNOLOGIES

Knowledge management refers to the acquisition, creation, dissemination and utilization of knowledge.[3] With knowledge management, businesses aim to increase efficiency, reduce costs and retain customer’s profile like purchasing behaviors, demographics includes age, gender, income, job title, educational level, family status etc, psychographics includes lifestyle, goals, habits, values, interest of customers, geographical factors includes city, area, region, country of customers etc. Knowledge management helps to better customer targeting. Companies can give direct larger discount offers to customers who are more likely to become repeat customers. Knowledge management reduces customer Acquisition Cost (CAC). CAC is the money which spends on marketing and sales campaigns to attract a single customer.[4] Knowledge management tools also help to reduce customer service expenses. It enables business to optimize hiring and training agents, who handles call monitoring and utilize integrated call center software. With this, Companies can interact with their customers and enable callback from the queue. For example, once you shopping from Naaptol, their customer service agents regularly call you to tell about new offers. Knowledge management considers four basic processes: knowledge creation, storage, transferring and applying. [5]

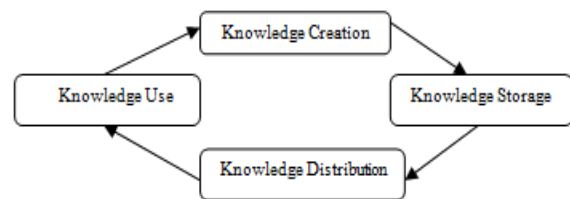


Fig. 2 Knowledge Management Process

Creation of knowledge takes place by the conversion of tacit and explicit knowledge and shown in Table 1 with together with examples of applied technologies. This conversion is represented by SECI (Socialization, Exteriorization, Combination, Interiorization) model.

Explicit Knowledge is also known as documented information. It is easy to express, write down and share. Tacit knowledge gained from personal experience that is more difficult to express or share with others.

Knowledge storage is a stage to store explicit knowledge. Knowledge is stored in the form of a knowledge repository, which includes databases, documents, reports etc. This repository allows many people to search for and retrieve particular knowledge without having to contact with the person who originally developed it. This saves time and improves performance of the organization. The most common storage technologies are CompactFlash, Secure Digital(SD), MemoryStick, MicroSD, Solid state drives etc.

Type of Knowledge	Tacit	Explicit
Tacit to	Socialization Sharing of experience through Video Conferencing, Phone Calls, Online Chat, Group interviews	Extetalization Creates conceptual knowledge through articulation using Expert Interviews, Ideation workshops
Explicit to	Internalization Creates operational knowledge through learning by doing. Manuals, Audio/Video stream Presentations	Combination Creates systematic knowledge through Search, Data Mining, records, Policies, Databases, Specifications

Table 1 Conversion of Tacit and Explicit Knowledge Knowledge Distribution refers to the process of sharing of explicit or tacit knowledge among individuals, groups and organizations. Knowledge sharing makes information available to those who actively look for it. Techniques used for knowledge sharing are audio-visual education, quizzes, conferences, coffee meetings, superior mentorship, team sharing etc. Information technology tools like email, databases, chat groups, discussion forums, video conferencing technologies, groupware, document management systems are integrated into knowledge sharing systems.

Knowledge applying refers to the use of available knowledge to make decisions and improve efficiency by reducing cost and solve business problems.[6] The knowledge application has become more comfortable these days with the help of technology. For example, several business intelligence tools like Artificial intelligence and Machine learning to process the knowledge data and offer analytics and reporting for better decision making. Document360 and HubSpot are popular knowledge management solutions that help businesses manage various steps of knowledge management process on a single platform.

#### IV. DATA MINING FOR KNOWLEDGE DISCOVERY IN E-COMMERCE

E-commerce is the buying and selling of goods, products or services over the internet. Transaction of money, funds and data are also considered as E-commerce. These business transactions can be done in four ways: B2B, B2C, C2C, C2B. Online stores like Amazon, Flipkart, Myntra, Ebay, Olx are some examples of E-Commerce websites. Because of the nature of the internet, ecommerce businesses need to obtain a lot of data about their customers. Data can be obtained whenever a purchase is made, an account is created or a page view is made. This data comes from the database associated with an ecommerce website as well as from web analytical tools like Google Analytics. Whenever a visit to a site is made, the information about the visitor is collected. This information could be about the visitor’s journey through the site, how they got on the site, what they searched to find the site, where they went after the site, how long they spent on each page of site etc. Data could also be about the visitors themselves. Companies can analyze customer’s behavior by their visit on a particular item again and again, purchase history, items in cart but not buying, items like etc. Data mining techniques helps companies to analyze the demographic of visitors to their site. The data mining techniques that are used in ecommerce are:

##### A. Association:

Association is well known data mining technique. The meaning of association is a temporal relation between attributes. In association, a pattern is searched based on a relationship of a particular item on other items in the same transaction.[7] Association

analysis is known as association rules. The best example of association rule is market basket analysis. It is a data mining technique used by retailers to increase sales by understanding the customer purchasing patterns such as purchase history, products that are likely to be purchased together.[8] It studies the frequency of items occurring together in databases and based on a threshold called support, identifies the frequent item sets. Another threshold, confidence, is the conditional probability of a transaction containing item set given that it contains other item set. In concerned with ecommerce, the shopping websites use market basket analysis to analyze customer's behavior. For example, those who buy diapers can also buy diaper rash cream together. Assume there are 100 customers. 10 of them bought diapers, 8 bought diaper rash cream and 6 bought both of them. Bought diaper bought diaper rash cream.

Support= $P(\text{diaper} \ \& \ \text{diaper rash cream}) = 6/100 = 0.06$ . Confidence= $\text{support}/P(\text{diaper rash cream}) = 0.06/0.8 = 0.75$ . lift= $\text{confidence}/P(\text{diaper}) = 0.75/0.10 = 7.5$ .

#### B. Clustering:

Clustering is the task of grouping a set of objects which have common behavior in the same class and the classes have unknown labels.[9] Clustering analysis is broadly used in market research, online advertisement and customer shopping behavior etc. Clustering can help sellers discover specific groups in their customer base and they can characterize their customer groups based on the purchasing pattern. The goal of cluster analysis in marketing is to segment customers in order to achieve more effective customer marketing.[10]

There are large numbers of customers who visit a particular site everyday. Companies can't connect with all the customers, but by dividing market into group of customers with similar behavior, companies can position themselves to appeal to these segments. Companies can identify homogeneous groups of buyers by their behavior on measures such as favourite stores, price willing to pay, favourite brand, frequency of purchase etc. for example, ecommerce data mining of a clothing website might show that 80 out of 100 visitors are female. 60 out of 100 aged 21-35. 92 out of 100 live in India. 20% of them are shown as "interest in sports". Here data mining

techniques help companies that it would be beneficial for them to begin to market and sell female sports clothing.

In the domain of online advertising, clustering helps to show a particular kind of ad to a group of users who visit a webpage on website to maximize the chance of a desired action to be performed after seeing advertisement. For example, users who visit the mobile website to purchase a mobile can also buy power bank by seeing advertisement.

#### C. Classification and Prediction:

Classification is a technique where data can be categorized into given number of classes according to its similarities.[11] This is the technique of arranging the data in different homogeneous groups. The main goal of classification is to identify the class to which new data will fall under. The algorithm used for classification in data mining is called the classifier, which is constructed to predict categorical labels. A classifier decides in which class an item belongs to on the basis of the values of its attributes. For example, Company gives discount to the customers on the basis of their shopping amount. If Shopping amount  $\geq 2500$ , then customers get 20% discount on their bill. If Shopping amount  $\geq 5000$ , then customers get 40% discount on their bill. If Shopping amount  $\geq 10,000$ , then customers get 60% discount on their bill. These discounts can attract customers to do more shopping. Prediction in business is done for the purpose to forecast or predict trends. It helps to make better decisions. In prediction, the model predicts continuous-valued function. This model is called predictor. The accuracy of a predictor is evaluated by computing an error based on the difference between the predicted value and the actual known value of y for each of the test tuples, x .[12] for example, Predictive data mining process use algorithms to predict future trends by going through customer's past transactions from database. The marketing manager can predict how much a given customer will spend during sale time.

The key feature of e-commerce is to provide global marketplace to consumers where they can buy products or services according to their needs anytime anywhere in the world.[13] So it is necessary to apply knowledge management applications and data mining techniques for enhancements in business, better

customer satisfaction, marketing of products or services etc.

#### V.CONCLUSION AND FUTURE WORK

In this paper, we discussed about data mining, knowledge management and how data mining is used for knowledge discovery in electronic commerce to improve the services provided by e-commerce based businesses. There are several techniques that have been developing and using in data mining processes in e-commerce. Data mining is useful to analyze a large amount of data for pattern discovery , predicting customers behavior, helps decision makers to make better decisions, reduce risk and helps in developing smarter marketing campaigns. Knowledge management aspect of an e-commerce system has also been discussed in this paper. To achieve better knowledge management, three fundamental functions must achieve. These functions are knowledge acquisition, knowledge transfer and application and knowledge use. At the end, we conclude that data mining for e-commerce businesses should no longer be privilege but requirement in order to survive and remain relevant in the competitive environment. The special issue highlights that although e-commerce systems are an ideal application for data mining. However, there still is must research needed in order to improve e-commerce marketing.

#### REFERENCE

- [1] N.R. SrinivasaRaghavan, "Data mining in e-commerce: A survey", April/June 2005.
- [2] Kamba, M. and Hang, J. (2006) Data Mining Concept and Techniques. Morgan Kaufmann Publishers, San Fransisco.
- [3] Pitt, M., Clarke, K.: Competing on Competence: A Knowledge Management Perspective on the management of Strategic Innovation. Technology Analysis and Stretegeic Mangement 12, 75-76 (2004).
- [4] Pei, X.: Enterprise Knowledge Management System in Knowledge Economy Condition. Management Engineering Journal 25, 26-28 (1999).
- [5] Rowley, J. (2000) Coupling e-business with knowledge management. Managing information June 59-62.
- [6] McAdam, R. and McCreedy, S. (1999) A critical review of knowledge management models. The Learning Organisation 6 (3): 91-101.
- [7] Agarwal R, Srikant R. "Fast algorithms for mining association rules", 1994.
- [8] Pankaj Gupta, Bharat Bhushan, " Data Mining Techniques in E-Business", International Journal of Advances in Engineering Sciences, Jan, 2012.
- [9] Barry, M.J.A. and Linoff, G.S. (2004) on Data Mining Techniques for Marketing, Sales and Customer Relationship Management. Indianapolis Publishing Inc., Indiana.
- [10] Pan, Q. (2011) Research of Data Mining Technology in Electronic Commerce. IEEE Computer Society, Wuhan, 12-14 August 2011, 1-4.
- [11] Turner, B.A. (1983) The use of grounded theory for the qualitative analysis of organizational behavior. Journal of Management Studies 20 (3): 333-348.
- [12] Jermstittiparsert, K., Pamornmast, C., &Sriyakul, T. 2014. "An Empirical Discourse Analysis on Correlations between Exchange Rate and Industrial Product Export." International Business Management 8 (5): 295-300.
- [13] Wigand, R.T. (1997) Electronic Commerce: definition, theory and context. The Information Society 13 (1): 1-15.