Role of Data mining in analyzing consumer’s online buying behavior

Bhumika Pahwa1, Dr. S. Taruna2, Dr. Neeti Kasliwal3
1(Information Technology, DAV Institute of Management, Faridabad, India)
2(Computer Science, JK Lakshmipat University, Jaipur, India)
3(School of Pharmaceutical Management, IIHMR University, Jaipur, India)

ABSTRACT: Online shopping is still in its nascent stage in India but growing at a fast pace. To continue its growth it is significant to understand the user’s preferences. Analysis of consumer’s behavior with respect to online shopping consists of detailed information about consumers past purchases as well as prediction of future purchases. This growing need for refined information can’t be met with simple database software. Data Mining is used for finding the hidden information from the pool of data. It has also been called as data analysis, knowledge discovery and deductive learning. The ability to recognize and track patterns in data help businesses sift through layers of seemingly unrelated data for meaningful relationships. Through this analysis it becomes easy for the online retailers to determine the dimensions that influence the uptake of online shopping and plan effective marketing strategies. This paper builds a roadmap for analyzing consumer’s online buying behavior with the help of data mining. The aim of this paper is to understand the role of data mining in growth of online shopping. The major factors that affect the consumer’s online buying behavior are convenience, ease of use and perceived benefits. Security is also a major consideration when opting to conduct shopping activities online. This study will help in further analyzing the consumer online buying behavior towards Online shopping which will help the retailers to design appropriate marketing strategies for selling their products online which will further help in development of the country.

Keywords - E-commerce, M-commerce, Data mining, Consumer Behavior, Personal perceived values, Website quality.

I. INTRODUCTION

E-commerce basically stands for electronic commerce which relates to a website that sells products or services directly from the site with the help of a shopping cart or shopping basket system and payments can be done through cards, e-banking and cash on delivery. It helps customers to buy anything form a pen to an insurance policy from the comfort of their home or office and gift it to someone sitting miles apart just by click of the mouse. It offers various benefits to businesses for example, easy reach to fast growing online community, providing unlimited shelf place for products and services, merging global markets at low operating costs. Ease to access internet is the major factor in rapid adoption to E-commerce. For popularization of E-commerce in India the main essential factors are safe and secure payment modes. Even though there are various benefits in shopping online but just like every coin has two sides, there exist various reasons for not shopping online for example lack of trust, security concerns, uncertainty about the product and service quality, delay or non-delivery of goods, and lack of touch-and-feel shopping experience.

Mobile Commerce (M-commerce) is the subset of electronic-commerce, which includes all e-commerce transactions carried out using a mobile device. Basically, M-commerce is the way of doing business in a state of motion. M-commerce depends on the availability of mobile connectivity. M-commerce offers multiple advantages like ubiquity, personalization, flexibility, and distribution, instant connectivity, immediacy. There are many ways in which businesses, government and people benefit from m-commerce like:

- Selling a product or service which is information based (delivery directly to mobile devices) or location based
- Improving productivity by gathering time critical information (reports, photographs) and SMS based up-to-date information.
- The ability to access information on mobile, at affordable cost can change people's lives and livelihoods in rural areas (Latest on the weather report or health services). It can be used as the medium to educate and create awareness among the rural people. Usages of Internet on mobile devices have lead to information access overcoming geographical barriers and removed the training cost of mobile technology.
Customer behavior analytics is based on consumer buying behavior, with the customer playing the roles of user, payer and buyer. The concern of many organizations is no longer on the individual buyer but rather on collective or organizational buying behavior which help in determining which customers are worth developing and managing by putting unique strategies in place in order to attract specific customers. Through analysis of customers’ behavior, accurate profiles are being generated by specifying needs and interest and allowing business to give customers what they want it, when they want, leading to a better customer satisfaction thereby keeping them to come back for more.

Consumer behavior includes the study of individuals, groups or organizations about their process of selecting, securing, using and disposing the products, services, experiences or ideas to satisfy needs and the impact of these processes on the consumer and society.

While large-scale information technology has been evolving separate transaction and analytical systems, data mining provides the link between the two. Data mining software analyzes relationships and patterns in stored transaction data based on open-ended user queries. Data mining is the semi-automatic discovery of patterns, associations, changes, anomalies, and statistically significant structures and events in data. Traditional data analysis is assumption driven in the sense that a hypothesis is formed and validated against the data. Data mining, in contrast, is data driven in the sense that patterns are automatically extracted from data. Various studies on consumer purchasing behavior have been presented and used in real problem. Data mining techniques are expected to be more effective tools for analyzing consumer behavior. Data mining has quickly emerged as a highly desirable tool for using current reporting capabilities to uncover and understand hidden patterns in vast database and these patterns are then used in models that predict individual behavior with high accuracy.

II. FACTORS THAT LEAD TO THE GROWTH OF ONLINE MARKET

A significantly low (19%) but fast-growing internet population of 243 million in 2014 is an indicator of the sector’s huge growth potential in India.

According to Forrester Research, an independent technology and market research firm, only 16% of India’s total population was online in 2013 and of the online users only 14% or 28 million were online buyers. India, therefore, was still in a nascent or immature stage of evolution of online retail spending. China was in ascending stage at 50%, whereas Japan (69%), Australia (57%) and South Korea (70%) were in mature stage.

According to Internet live stats website, it is evident that in absolute terms India’s internet users are short by only 36 million as compared with 279 million in the US and higher than that in Japan, Brazil and Russia. However, in relation with its population, only 19% Indians use the internet. This indicates the potential of internet use in India and as internet penetration increases, the potential of growth for the e commerce industry will also increase.

<table>
<thead>
<tr>
<th>Internet users by country: in million (2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
</tr>
<tr>
<td>US</td>
</tr>
<tr>
<td>India</td>
</tr>
<tr>
<td>Japan</td>
</tr>
<tr>
<td>Brazil</td>
</tr>
<tr>
<td>South Korea</td>
</tr>
</tbody>
</table>

Source: Internet Live Stats website accessed on 9 December 2014

Figure 1: Country-wise Internet Users

An analysis of the demographic profile of internet users further testifies that e-Commerce will rise rapidly in India in coming years. Around 75% of Indian internet users are in the age group of 15 to 34 years. This category shops more than the remaining population. Peer pressure, rising aspirations with career growth, fashion and trends encourage this segment to shop more than any other category and India, therefore, clearly enjoys a demographic dividend that favors the growth of e-Commerce. In coming years, as internet presence increases in rural areas, rural India will yield more e-Commerce business.
Role of Data mining in analyzing consumer’s online buying behavior

III. M.COMMERCE

Anshu et.al, 2015 stated the reasons why m-commerce is growing in India like, low price of the devices, easy availability of internet on mobile phones and 24x7 portability and availability of the device, and growth barriers as security, ROPO(Research online and purchase offline tendency of people), language barriers, acceptance of new technology, complex mobile applications [1].

Paul Monno & Dazhi Xiao, 2014 [2] investigated the reasons why consumer adoption behavior of m-commerce apps in the European market differs from those in the Asian market and found that European consumers lacked knowledge and were unable to perceive the full conveniences of using apps, compared to Asian respondents and that European consumers placed a lot of importance on risks and anxieties when adopting apps.

Gupta & Vyas, 2014 [3] says that m commerce is at emerging level in India and it is complex to adapt. The benefits are user-friendly, easy to carry; low internet connective area etc. and drawbacks are lack of internet connectivity, language barrier, less graphic resolution, less number of mobile phone users etc.

Kanwalvir Singh, 2013 [4] analyzes and measures the factors influencing the consumer’s attitude towards m-commerce and mobile services and determines what users think about the various existing m-commerce practices already available. The outcomes of his research were the critical success factors influencing the user’s intention to adopt M-commerce services as:

- Respondents under 18 years of age considered advantages of M-commerce in offering various types of M-transaction services & also due to the affordable price of transactions.
- Respondents in the urban region considered M-commerce to be more flexible & having affordable price. Respondents in the age group of 18-60 years & residing in rural regions have been found to be more concerned towards hindrances in consumer adoption of M-commerce due to the existing digital divide between urban and rural citizens.
- Majority of the respondents in the age group of 18-60 years & staying in rural regions found M-commerce difficult to use because of the less no. of training and guidance facilities for the rural citizens. Other factor might be the language barriers or constraints existing within the rural regions.

Niranjanamurthy et al. 2013 [5] identify several disadvantages of m-commerce: firstly increased security risks related to the less sophisticated operating systems on mobile devices, as well as accessibility issues, due to small screen size.

Batra & Juneja, 2013 [6] has extended their research on mobile commerce in India. It lists the issues faced by M-Commerce Industry. He had also studied the sales and available users of smart phones in India which shows tremendous increase in the graph although it is new in India. According to him the growth drivers of m-commerce are instant connectivity, personalization factor, mobility factor, immediacy, localization etc.

Mirzae, Asadollahi, & Jahanshahi, 2012 [7] have proposed a study about brief description about mobile wireless technologies, relationship between e commerce and m commerce, help business to define what they can derive from m commerce. M-commerce can be used for Travel and Ticketing, Education, Health care, Traffic etc. he had also brought them into consideration about the use of m commerce in China and Brazil and India, about how it has increased from 2010 to 2011. At last they also suggested m-commerce providers to improve user interface, and implement innovative pricing structures.

Amin Asadollahi et al, 2011[8], stated that Mobile commerce is a natural successor to electronic commerce and M-commerce should not be viewed as E-commerce with limitations, but rather as a unique form of e-commerce with its own unique benefits and ubiquity, intimacy, time sensitivity and location awareness are key concepts that make mobile commerce so different from traditional e-commerce.
IV. CONSUMER BEHAVIOR

Consumer behavior means the study of individuals, groups or organizations about their process of selecting, securing, using and disposing the products, services, experiences or ideas to satisfy needs and the impact of these processes on the consumer and the society. Behavior concerns either with the individual or the group (e.g. In college friends influence what kind of clothes a person should wants to wears) or a firm (peoples working in firm make decision as to which products the firm should use.) The use of product is often so important to the marketer because this may influence how a product is best positioned or how they can encourage increased consumption. Consumer behavior involves services and ideas as well as tangible products.

Donal Rogan, 2007[9] explained the relationship between consumer behavior and marketing strategy. He has stated that “strategy is about increasing the probability and frequency of buyer behavior. Requirements for succeeding in doing this are to know the consumer and understand the consumer’s needs and wants.”

Chisnall, 1995[10] points out that human needs and motives are inextricably linked and that the relationship between them is so very close that it becomes difficult to identify the precise difference which may characterize them. People may buy new jackets because it protects them against the weather, but their real underlying dominant need may be to follow the latest fashion trend.

Kotler and Armstrong, 2007 [11] explain the ways in which the consumer interprets and receives stimuli from advertisements. The decisions of consumers are influenced by a number of individual characteristics that are linked to the consumer’s specific needs.

Table 1 explains the factors that affect the consumer behavior.

OBJECTIVE

The major objective of this paper is to focus on topic that how Data mining can be used as an analysis tool for analyzing the consumer’s attitude towards e-commerce and m-commerce. This paper outlines the process of how data mining works. It provides a framework for the analysis. In this paper the research gaps in the field of consumer behavior with respect to data mining. This will further help the researchers to gain directions in this field.

DATA MINING

Data mining is one of the most powerful new technologies with great potential to help companies focus on the most important information in the data they have collected about the behavior of their customers and potential customers. It discovers information within the data that queries and reports cannot effectively reveal. It automates the detection of relevant patterns in a database. [12, 13,14,15,16,17,18,19]

<table>
<thead>
<tr>
<th>References</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>San José Cabezudo, 2010</td>
<td>Personality traits</td>
</tr>
<tr>
<td>Hill and Beatty, 2011; Hernández, Jiménez and Martin, 2011</td>
<td>Self-efficacy</td>
</tr>
<tr>
<td>San José Cabezudo, 2010</td>
<td>Demographic profiles</td>
</tr>
<tr>
<td>Close and Kukar – Kinney, 2010</td>
<td>Acceptance of new IT applications</td>
</tr>
<tr>
<td>San José Cabezudo, 2010</td>
<td>Perceived danger</td>
</tr>
<tr>
<td>Lim and Dubinsky, 2004</td>
<td>Perceived convenience</td>
</tr>
<tr>
<td>O’Cass and Fenech, 2003</td>
<td>Perceived website quality</td>
</tr>
<tr>
<td>Hernández, Jiménez and Martin, 2010</td>
<td>Perceived benefit</td>
</tr>
<tr>
<td>Chaipoopirutana and Combs, 2010</td>
<td>Security</td>
</tr>
<tr>
<td>Galanxhi-Janaqi and Fui-Hoon Nah, 2004</td>
<td>Privacy</td>
</tr>
</tbody>
</table>

For example, a pattern might indicate that married males with children are twice more likely to drive a particular sport car than married males with no children. For a marketing manager of an auto mobile industry, this surprising pattern might be valuable.

According to [20] Jiawei Han et. al. 2006, Data Mining is the semi-automatic discovery of patterns, associations, changes, anomalies, and statistically significant structures and events in data. Traditional data analysis is assumption driven in the sense that a hypothesis is formed and validated against the data. Data mining, in contrast, is data driven in the sense that patterns are automatically extracted from data (Jiawei Han et.al. 2006).

Data Mining is an iterative and interactive discovery process. Its goal is to mine patterns, associations, changes, anomalies, and statistically significant structures from large amount of data. Further-more, the mined results should be valid, novel, useful, and understandable.

According to Witten and Frank, 2005 [21] data mining means “solving problems by analyzing data that already exists in databases”. Mitchell, 1997 [22] wrote “Machine Learning is a mature and well-recognized research area of computer science, mainly concerned with the discovery of models, patterns, and other
regularities in data”. Data mining generally refers to “extraction of hidden information from the large amount of databases.” Dunham et al. 2003 [23] cite “Data mining helps organizations to discover and identify the hidden patterns in databases. The extracted patterns are then made in use to build data mining models and hence can be used to predict performance and behavior with high accuracy.”

Lot of work has been done in previous years in the field of data mining with respect to consumer behavior analysis:

Junzo Watada et al, 2006 [24] tried to improve data mining analysis by applying several methods including fuzzy clustering, principal component analysis and discriminate analysis in their paper ‘A Data mining approach to consumer behavior’.


Shrivastava A et al, 2007[26] explain that Data mining is an attitude that business actions should be based on learning, that informed decisions are better than uninformed decisions, and that measuring results is beneficial to the business. Data mining is also a process and a methodology for applying the tools and techniques. Association rule mining is also one among the most commonly used techniques in Data mining. A typical and the most running example of association rule mining is market basket analysis.

DATA MINING AND CONSUMER BEHAVIOR

Information system can query past data up to and including current level of business. Often businesses need to make strategic decisions or implement new policies that better serve their customers. For example, Grocery store redesigns their layout to promote more impulse purchasing. Telephone companies establish new price structures to entice customers into placing more calls. Both task require an understanding of past customers consumption behavior data in order to identify pattern for making those strategic decisions and data mining is particularly suited to this purpose. With the application of advanced algorithms, data mining uncovers knowledge in a vast amount of data and point out possible relationships among the data. The core components of data mining technology have been developing for decades in research areas such as statistics, artificial intelligence and machine learning. Today technology is mature and when coupled with relational database systems and a culture of data integration they create a business environment that can capitalize on knowledge formally buried within the systems [27].

DATA MINING AND DATA EXTRACTION

With the Advancement in technology, various changes in approaches to organizing and retrieving information have been noticed taking advantage of the available data. Data mining gives us the ability to see patterns, predict the future and make informed decisions based on the evidence in large databases. For example, data mining of categorical and numerical consumer shopping data allow retailer to understand which items are purchased by the same customers, predict sales of seasonal items and more efficiently manage its inventory [27]. Primarily the data mining requires a standard process, data store or warehouse, technologies and expertise. The process must be reliable and repeatable by people with little data mining skills. However the standard data extraction process should involve job understanding which determines the job objectives, job background situation assessment etc., followed by the data understanding which collects data, describes data, explore data, and verify data quality. The preparation involves the data set description, selection, assessment, consolidation, data formatting etc. process modeling, process evaluation and deployment [27].

STAGES IN DATA MINING

Data mining is an analytic process designed to explore data (usually large amounts of data) in search of consistent patterns and or systematic relationship between variables and then to validate the findings by applying the detected patterns to the new subset. Data mining is often considered as a blend of artificial intelligence and statistics [27].

The process of data mining consists of three stages which are:

- **The initial exploration**: This stage usually begins with data preparation which may involve cleaning data, data transformation, selecting subsets of records. This first stage of data mining may involve anywhere between a simple choice of straight forward predictors for a regression model to elaborate exploratory analyses using wide variety of graphical and statistical methods in order to identify most relevant variables and determine the complexity and or the general nature of models that can be take into account in the next stage.

- **Model building or pattern identification**: This stage involves considering various models and choosing the best one based on their predictive performance i.e. explaining the variability in questions and producing stable results across samples, this may sound simple but it involves an elaborate process. There are several
techniques that can be applied to achieve that goal many of which are based on applying different models to the same data set and comparing their performances to choose the best.

- Deployment: This involves using the model selected as the best in the previous stage and applying it to the data in order to generate predictions or estimate the expected outcome.

Classification and prediction in data mining

There are two forms of data analysis that can be used for extracting models describing important classes or to predict future data trends. These two forms are Classification and Prediction. (Jiawei han, ‘Data mining concepts and techniques’) Classification models predict categorical class labels; and prediction models predict continuous valued functions. For example, a bank loan officer wants to analyze the data in order to know which customer loan applicants are risky or which are safe. Similarly here, the motive is to find which factor’s effect on the consumer behavior for shopping online using m-commerce with respect to e-commerce are positive or negative. Classifiers will be constructed to predict the categorical labels; these labels are positive or negative for the consumer’s attitude.

The Data Classification process includes two steps −

- Building the Classifier or Model
- Using Classifier for Classification

Building the Classifier or Model

This step is the learning step or the learning phase. In this step the classification algorithms build the classifier. The classifier is built from the training set made up of database tuples and their associated class labels. Each tuple that constitutes the training set is referred to as a category or class. These tuples can also be referred to as sample, object or data points.

Using Classifier for Classification

In this step, the classifier is used for classification. Here the test data is used to estimate the accuracy of classification rules. The classification rules can be applied to the new data tuples if the accuracy is considered acceptable.

IMPORTANCE OF DATA MINING IN MODERN BUSINESS

Data mining is used in reestablishment of hidden information of the data of the algorithms. It helps to extract the useful information starting from the data, which can be useful to make practical interpretations for the decision making. Data mining can be technically defined as automated extraction of hidden information of great databases for the predictive analysis. In other words, it is the retrieval of useful information from large masses of data, which is also presented in an analyzed form for specific decision making. Although data mining is a relatively new term, the technology is not. Data mining is thus also known as Knowledge discovery in databases since it grip searching for implied information in large databases. However, the use of some advanced technologies makes it a decision making tool as well. Data mining is used in market research, industry research and for competitor analysis. It has applications in major industries like direct marketing, ecommerce, customer relationship management, scientific tests, genetics, financial services and utilities. Data mining applications are often structured around the specific needs of an industry sector or even tailored and built for a single organization. This is because the patterns within data may be very specific. The value of data mining applications in business is often estimated to be extremely high. Some businesses have stored large amounts of data over years of operation, yet without an appropriate data mining application are missing out on the very valuable information that may be contained within their existing data.

V. CONCLUSION

Data mining is primarily used today by companies with a strong customer focus - retail, financial, communication and marketing organizations. Data mining has a lot of importance because of its huge applicability. It is being used increasingly in business applications for understanding and then predicting valuable data, like customer buying actions and buying tendency, profiles of customers, industry analysis, etc. Data Mining is used in several applications like market research, customer behavior, direct marketing, bioinformatics, genetics, text analysis, e-commerce, customer relationship management and financial services. Studying consumer Online buying behavior with the help of data mining as the analysis tool will give a better insight about the perception of the users when they shop online. This will further help the online retailers to plan effective marketing strategies, which will lead to growth in the online market that will generate profits and further will lead to growth of the Indian economy.
REFERENCES


[10] Peter M. Chisnall, Strategic business marketing (Prentice Hall, 1995)


International Journal of Business and Management Invention (IJBMI) is UGC approved Journal with Sl. No. 4485, Journal no. 46889.