INTERPRETING CUSTOMER BEHAVIOUR PATTERN USING MINING TECHNIQUES

ANU DAHIYA*

Former Assistant Professor at Hindu Institute, Sonepat, India.

ABSTRACT

This paper identifies the issues of customer relationship management through customer segmentation and customer profiling .The text also examines the ways of applying data mining techniques to identify association between customers behaviour and their buying patterns. Customer's buying pattern will be identified using segmentation technique and their behaviour will be identified using customer profiling. For one segment K-Means clustering algorithm has been used and for other part make density based clusters are employed. The purpose of identifying these traits of customer is to investigate how these behaviour will affect the sales and revenues in business.

1. INTRODUCTION

Nowadays everything is being recognized as universal market that has become must have component for organization success. It has nearly get difficult to all contender to hold out in market for a sustained life because competition is hard. Variety is the soul faith of market. This is how development of correct business strategy over time is developed. Appropriate business strategy is something which helps companies achieves business objectives. Business objectives help achieve corporate objectives and corporate objectives aim to achieve a competitive advantage over rival organizations. Efficient business strategies or business campaigns often consist of a collection of various business tactics that work together in a interactive way to base your brand, minimize sales immunity, and develop interest and hope for our items and services. Now, business is all over, conversationally or formally, people and organization mesh in large number of activity that we describe as business.

E-business Customer visibilities Research studies indicate that the E-Business environment is competitive and opportunities for ecommerce success exist; Nevertheless, Unsuccessful business does exists. E-business has activated businesses to expand to a world market which in turn enhance the customer's foundation, sales and accordingly profits. Online businesses have to confront a vast amount of challenges, one very important challenge being to identify online customers' choices. Numerous techniques has been developed as resolution, for instance, one-to-one online business strategies. The techniques assist E-Business businesses to survive on the Internet by incrementing customer's commitment and bending E-Business website visitors into customers. Progressively online market has discovered the important part customer profiles play in discovering online customers' demands and preferences. A customer profile is a snapshot of our customers, how to reach them and why they buy from us. Briefly, a customer profile is a combination of data that differentiates the customer. Customer profiling is the process of making a profile using applicable and usable information to describe the features of an isolated customer and to discover differentiators from other customers and drivers for their buying decision. Customer profiles can be differentiating by two sets of information namely actual or stable and behavioural or active. Actual or static profile information contains specific facts about the customer, including demographics, for example, age, gender and name. Active profile information demonstrates the behaviour of the customer and this is done using continuative rules such as association and classification rules. An instance of behaviour is when buying on weekends; Customer Y usually spends R1000 on items. By giving customer profiles, businesses can filter and discover information to help manipulate and make efforts to meet the demands and tastes of each unique customer or a group of customers based on their profiles. Customer profiles provide a suitable tool to analyse and determine customers' needs and preferences that can be used to implement one-on-one online business schemes. Some famous online businesses that use customer profiles include alahari.net in South East and StalkBuy.com in the United Kingdom. Explicit and implicit results are two basic ways utilized by online businesses to establish a customer profile. Explicit feedback is the simplest method to establish a customer profile. Customers are openly asked to do their entry on the website using an online questionnaire. Information that is recorded during this process normally includes actual or stable information, for instance, full name, gender, customer's age and other demographic details. In some examples the registration continues by asking the customer to provide preferences, answer a specific questionnaire or rate items or products on the website. These sections for registration include behavioural or dynamic information that is used to model the users' online behaviour.

*Corresponding Author: Anu Dahiya**
Former Assistant Professor at Hindu Institute, Sonepat, India.

The paper analyzes behaviour patterns of customers through segmentation and profiling techniques. Segmentation includes clustering the dataset in to different segment according to the shopping behaviour of customer and profiling includes generating profiles of different customers according to different attributes for example: age, gender, income, country method of shopping etc.

The Text has been divided into following sub sections. Each section contains the following contents:-

Section - 2: Segmentation and Profiling
Section - 3: Conceptual Framework
Section - 4: Segmentation of Customers
Section - 5: Profiling Of customers
Section - 6: Conclusion

2. SEGMENTATION AND PROFILING

Market segmentation is a marketing strategy which involves dividing a broad target market into subsets of consumers, businesses, or countries who have, or are perceived to have, common needs, interests, and priorities, and then designing and implementing strategies to target them. Market segmentation strategies are generally used to identify and further define the target customers, and provide supporting data for marketing plan elements such as positioning to achieve certain marketing plan objectives. Businesses may develop product differentiation strategies, or an undifferentiated approach, involving specific products or product lines depending on the specific demand and attributes of the target segment.

Segmentation identifies how distinct profiling attributes inter-relate to make valuable structures and clusters in between your clients. Sometimes the segments can be as simple as different age bands or different geographic regions, but sometimes the analysis tools we provide can help you to identify patterns in behavior via attributes that simply aren't visible to the naked eye, allowing you to adjust your marketing strategies to achieve the optimum effect.

Combined together, Profiling and Segmentation exercises can frequently lead you towards a better understanding of why certain areas of your business may be performing poorly, or why your latest marketing campaigns may have suffered from a low response rate.

So how can you create an effective business strategy, deliver an impressive return on marketing investment and grow your market share when the likelihood is that you currently know very little about your clients? To achieve all of this and more, you need to have access to a marketing analytics solution which contains constantly refreshed, accurate transactional data that is serviced by a clean, unified and relevant data repository.

Customer Segmentation Strategy

Consumer market can be clustered on the following customer's traits:
• Geographic
• Demographic
• Psychographic

Behavioural Geographic Segmentation: Behavioral segmentation divides consumers into groups according to their knowledge of, attitude towards, usage rate, response, loyalty status, and readiness stage [6] to a product. There is an extra connectivity with all other market related sources. Behavioral segmentation divides buyers into segments based on their knowledge, attitudes, uses, or responses concerning a product. Many marketers believe that behavior variables are the best starting point for building market segments.

Demographic Segmentation: Segmentation according to demography is based on variables such as age, gender, occupation and education level or according to perceived benefits which a product or service may provide. Benefits may be perceived differently depending on a consumer's stage in the life cycle. Demographic segmentation divides markets into different life stage groups and allows for messages to be tailored accordingly.

Psychographic Segmentation: Psychographic segmentation, which is sometimes called lifestyle, is measured by studying the activities, interests, and opinions (AIOs) of customers. It considers how people spend their leisure,[8] and which external influences they are most responsive to and influenced by. Psychographics are very important to segmentation, because psychographics identify the personal activities and targeted lifestyle the target subject endures, or the image they are attempting to project. Mass media has a predominant influence and effect on psychographic segmentation. Lifestyle products may pertain to high involvement products and purchase decisions, to speciality or luxury products and purchase decisions.

Attitudes Behavioural Segmentation: Behavioural Segmentation is based on actual customer behaviour towards products. Some behavioural variables include:

• Benefits Sought
• Usage Rate
• Brand Loyalty

Profiling

Profiling is all about building up reservoirs of knowledge about your most typical clients. First, you need to understand who your clients are (e.g. their age, location, life-stage, income band, property value and lifestyle choices). Next, you need to build an understanding of how they are interacting with your products and services (e.g. How often do they purchase, how much, and when). Finally, you need to establish the way clients currently perceive your products and services (for example, whether you're providing a necessity, a luxury or a default option).

3. CONCEPTUAL FRAMEWORK

For solving the problem of identification of suspicious and non-suspicious transaction we have follow up the following procedure.

Data Mining And Segmentation Methods

Data Mining automates the detection of relevant patterns in a database, using defined approaches and algorithms to look into current and historical data that can then be analyzed to predict future trends. Because data mining tools predict future trends and behaviors by reading through databases for hidden patterns, they allow organizations to make proactive, knowledge-driven decisions and answer questions that were previously too time-consuming to resolve.

Data mining is not particularly new — statisticians have used similar manual approaches to review data and provide business projections for many years. Changes in data mining techniques, however, have enabled organizations to collect, analyze, and access data in new ways. The first change occurred in the area of basic data collection. Before companies made the transition from ledgers and other paper-based records to computer-based systems, managers had to wait for staff to put the pieces together to know how well the business was performing or how current performance periods compared with previous periods. As companies started collecting and saving basic data in computers, they were able to start answering detailed questions quicker and with more ease.

Segmentation

Segmentation or Clustering is a data mining technique that makes meaningful or useful cluster of objects which have similar characteristics using automatic technique. The clustering technique defines the classes and puts objects in each class, while in the classification techniques, objects are assigned into predefined classes. To make the concept clearer, we can take book management in library as an example. In a library, there is a wide range of books in various topics available. The challenge is how to keep those books in a way that readers can take several books in a particular topic without hassle. By using clustering technique, we can keep books that have some kinds of similarities in one cluster or one shelf and label it with a meaningful name. If readers want to grab books in that topic, they would only have to go to that shelf instead of looking for entire library. There are several techniques of data mining such as:

1. **Artificial neural networks:** Non-linear predictive models that learn through training and resemble biological neural networks in structure.
2. **Genetic algorithms:** Optimization techniques that use processes such as genetic combination, mutation, and natural selection in a design based on the concepts of natural evolution.
3. **Decision trees:** Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classification of a dataset.
4. **Nearest neighbour method:** A technique that classifies each record in a dataset based on a combination of the classes of the k record(s) most similar to it in a historical dataset (where k 1). Sometimes called the k-nearest neighbour technique.
5. **Rule induction:** The extraction of useful if-then rules from data based on statistical significance.
6. **Data visualization:** The visual interpretation of complex relationships in multidimensional data. Graphics tools are used to illustrate data relationships.

The approach in this project is using - **K MEAN** and WEKA tool for clustering and revenue profiling of data.

K-MEAN: K-Means Clustering is a method of vector quantization originally from signal processing, that is popular for cluster analysis in data mining. k-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as prototype of the cluster. This results in a partitioning of the data space into Voronoi cells. The problem is computationally difficult (NP-hard); however, there are efficient heuristic algorithms that are commonly employed and converge quickly to a local optimum. These are usually similar to the expectation-maximization algorithm for mixtures of Gaussia distributions via an iterative refinement approach employed by both algorithms. Additionally, they both use cluster centres to model the data; however, k-means clustering tends to find clusters of comparable spatial extent, while the expectation-maximization mechanism allows clusters to have different shapes.

Weka Tool

Weka (Waikato Environment for Knowledge Analysis) is a popular suite of machine learning software written in Java, developed at the University of Waikato, New Zealand. Weka is free software available under the GNU General Public License. The Weka workbench contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to this functionality.

Weka is a collection of machine learning algorithms for solving real-world data mining problems. It is written in Java and runs on almost any platform. The algorithms can either be applied directly to a dataset or called from your own Java code.

Figure 4.1: Thumbnail of Weka

4. SEGMENTATION OF CUSTOMERS

The data set of about 200 customers has been simulated for solving this problem. Firstly the attributes has been identified:
Cid
Cname
Age
Gender
Salary
City
State
Country
Method of shopping
Frequency
Item1
Item2
Item3
Item4
Net
Where Item 1 represents Food products
Item 2 represents Grossery things,
Item 3 represents toys
And item 4 represents accessories.
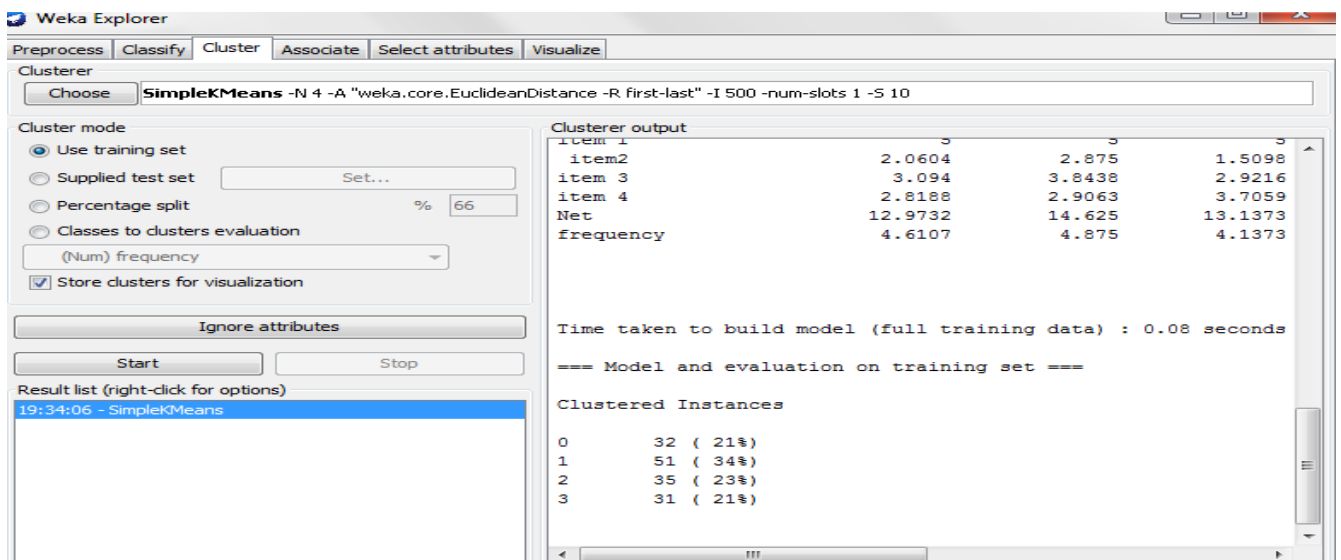
Here is the segmented behaviour pattern of customer:

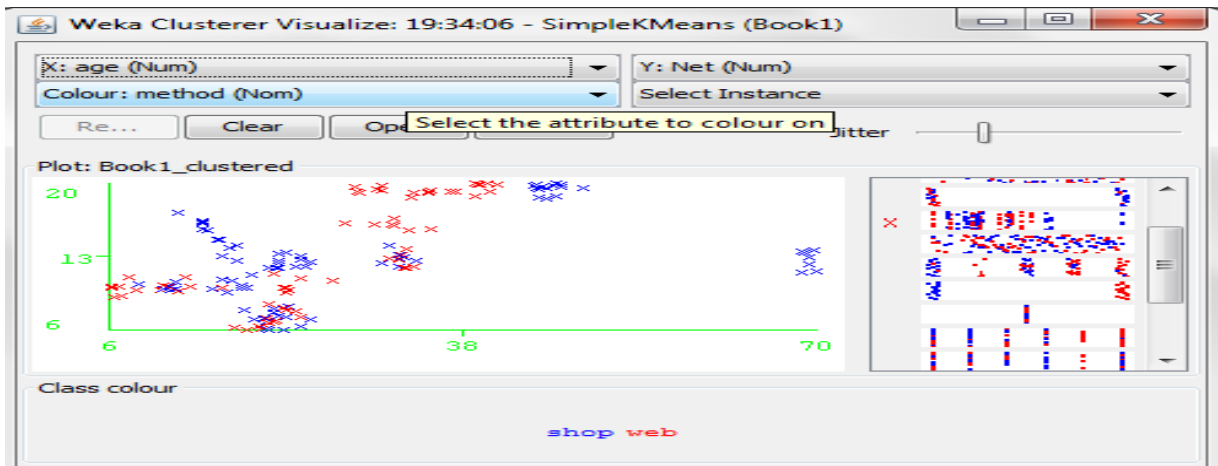**Figure 4.1:** Segmented Instances of customer data base

Figure 4.2: Clusters divided according to method of shopping

Transaction Id	Age	Net	Gender	frequency	Monthly Income($)	Segments
1	26.0625	14.625	Female	4.875	88375	Segment 0
2	25.0196	13.1373	Male	4.1373	89843.1373	Segment 1
3	24.2857	12.2	Female	5.2	91028.5714	Segment 2
4	34.3871	11.871	Male	4.4516	87161.2903	Segment 3

Table 4.1: Shows distribution of customers based on annual income, age and frequency.

From above description of data set, four segments has been analyzed for their behaviour. The very 1st segment which contains female customers has less tendency to shop than other female customers which are having average age group of 24 and annual income about $ 91028.5714.Also their Net shopped items consists of 12.2 and frequency 5.2. If segments are analyzed according to frequency of shopping than the segment 2 is more confident than that of 1st one.

If male clusters are analyzed, than segment 3 is more confident than segment 1 because frequency of latter is higher.

5. PROFILING OF CUSTOMERS

Customers are grouped using two attributes: - Age and Method of shopping i.e. whether they shop by web or by going to shop physically. When age has been used as a profiling parameter then two clusters are formed:-
- ✓ Cluster 0 contains total of 67 entries having average age value 25. And salary 89176 also the customers in this category are females.
- ✓ Cluster 1 contains total of 85 entries having average age value 28 .And salaries 83308 and customer in this category belong male group.

Which shows that the male customers having age about 28 (acc. to dataset) are more regular customer than that of female and having lesser salary and age. That mean the customers lies in second group are more confident customers than others.

Now this analyzed data can be used to identify more regular and confident customer by company to increase in their sales and final revenues.
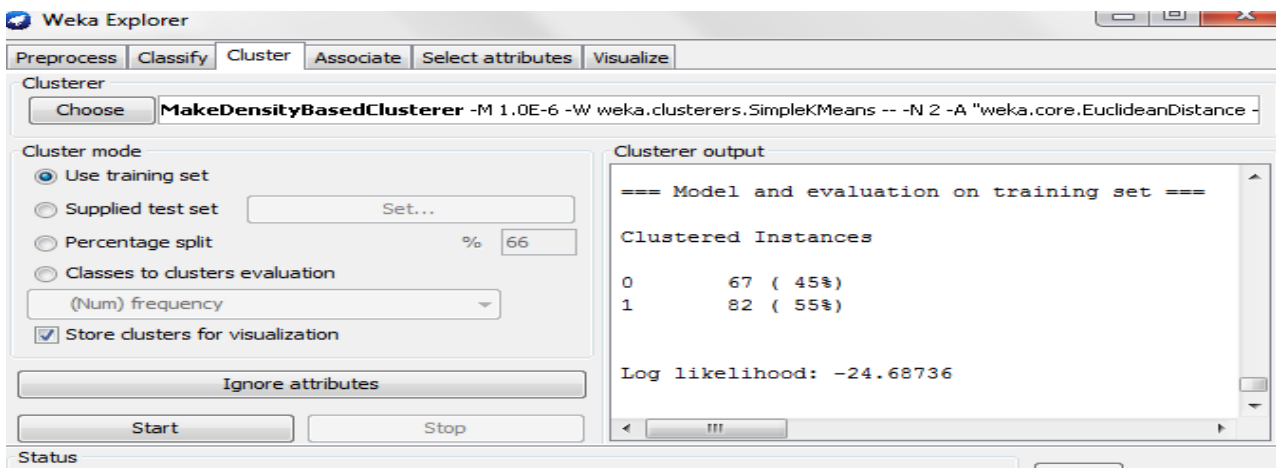

Figure 5.1: Profiling of customers according to their age

5. CONCLUSION

Generally, there is authorization of statistic elements for customer segmentation and profiling but severely, It has been analyzed that there is huge impact of orthogonal variables such as age, price, annual income, address and business trends on the buying by the consumers. Moreover, the intensity of connection between conventional basis and consumer segmentation depends mainly on many more orthogonal variables catered by business and consumer's conditions; that is why, work in this area would add some value to the literature in the field of segmentation. It is also discovered that segmentation is completely counted upon all conventional bases not on single one. This indicates that there is demand of further research in customer segmentation area in different specified areas to find the dominating basis of customer segmentation.

REFERENCES

1. https://www.coursehero.com/file/12693710/647-1598-1-SMpdf/
2. Canfora, G., Fasolino, A. R., Frattolillo, G. and Tramontana, P. (2008): A rapping approach for migrating legacy system interactive functionalities to Service Oriented Architecture. Journal of Systems and Software 81(2008): pp. 463-480.
3. Chung, R., Sundaram, D. and Srinivasan, A. (2007): Integrated personal recommender systems. ICEC. Minneapolis, Minnesota, USA. August 2007.
4. Cingil, I., Dogac, A. and Azgin, A. (2000): A broader approach to personalization. Communication of the ACM 43(8): pp. 136-141. August 2000.
5. Computer Economics (2009): SOA Adoption Surges. Available online: http://www.computereconomics.com/ article.cfm?id=1423&tag-rbspot. Accessed March 2009.
7. Ebay.com (2008): http://www.ebay.com. Accessed August 2008.
8. Eirinaki, M. and Vazirgiannis, M. (2003): Web mining for web Personalisation. ACM transactions on Internet technology 3(1): pp 1-27. February 2003.
9. Erl, T. (2005): Service-Oriented Architecture: concepts, technology and design. Prentice hall, Upper Saddle River, NJ, USA.
10. Gartner (2009): Press release: Gartner says SOA is evolving beyond its traditional roots. Available online: http://www.gartner.com/it/page.jsp?id=927612. Accessed May 2009.
11 Victoria K. Wells, Shing Wan Chang, Jorge Oliveira-Castro & John Pallister (2010), "Market Segmentation from a Behavioral Perspective", Journal of Organizational Behavior Management.
12 https://en.wikipedia.org/wiki/Market_segmentation.

Source of support: Nil, Conflict of interest: None Declared