National Conference on Advancements in Computer Science and Engineering
In association with International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)

Customer Segmentation Analysis

Mr. Manjunatha Swamy[1], Payel Pattanayak[2], S. Jagan[2], Mohith Teppola[2]

[1]Senior Assistant Professor, Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, Karnataka, India

[2]Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, Karnataka, India

ABSTRACT

Customer Segmentation Analysis is one of the business growth support systems. This is used to grow the business in the competitive market. This paper proposes an approach to develop a web portal based on target customers using predictive approach. After identifying the customers of same taste, they will be served products accordingly.

Keywords : Clustering, Customer Segmentation, Data Mining, Personalization, Ecommerce.

I. INTRODUCTION

With the evolution of new technologies and increasing growth of e-commerce it is important for every business to adapt new strategies which help them to win the competitive environment. Customer is the most valuable asset for any business. In this emerging market it is very difficult to maintain its customer base. To overcome this difficulty in a business has to focus on customer segmentation.

Customer segmentation means categorizing all the customers into same group. It is a technique in which we divide the customers based on their purchase history, gender, age, interest, etc. It is required to get this information so that the store can get help in personalize marketing and provide customers with proper deals. With the help of this project, companies can run targeted campaigns and provide user-specific offers rather than providing same offer to everyone. It acts as a base for Customer Relationship Management (CRM) for businesses to know the customer according to their transaction and focus on them separately. There are already various existing predictive models which provide info on customer segmentation and which helps them to segregate customers. For a successful business making more profit from each customer is key task apart from retaining and adding new customers. Many Different variety of models exist which helps the business managers to implement the strategic ways according to individual customer taste but each one with its own limitations.

II. RELATED WORK

Identifying right customer and providing right service at right time and treating different kind of customers differently is the key to run the business successfully.

So, a predictive model will be used to distinguish customers into different groups based on the data collected. Once the customers are segregated then their behavioral buying pattern are noted to increase the profit for the organization future new customer.

Integrated Approach of finding customer segments along with their associative buying pattern is as shown below

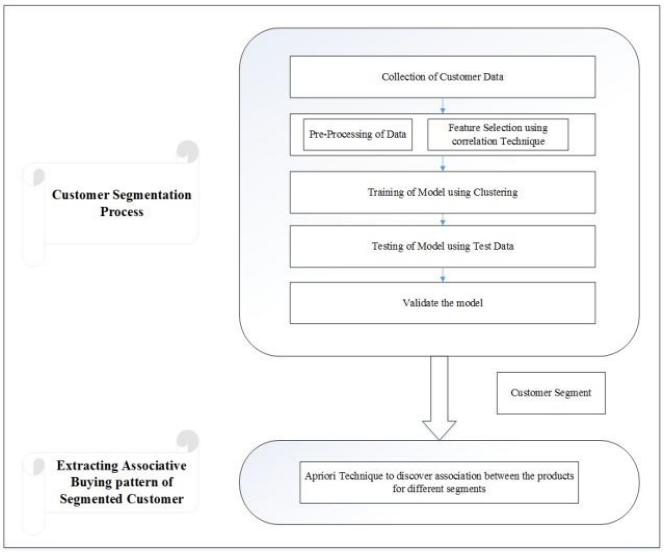

Fig. 1: Integrated Approach of finding customer segments along with their hidden buying patterns

Phase 1: Customer Segmentation

- **Step 1:** Collecting the Customer Data: It involves in the collecting the customer's transactional data that comprises of their statics (E.g.: Age, Gender etc.) and dynamic data (Eg: Purchase frequency etc.) from shopping vendors.
- **Step 2:** Pre-processing of Data: Pre processing of the data is one of the important steps for the accuracy of predictive model. In this step, the collected data will be segregated and required features will be extracted. Feature selection is responsible for extracting relevant features required for input vector of predictive model; it is a data reduction technique. This helps in as creating a subset of original features by not including the features which are redundant.

This paper proposes correlation technique helps in extracting the relevant features of the customers. Correlation measures the relationship between two features. It basically filters those features which are not required to form subset of the original features. To measure the relation between the two features, the correlation coefficient is taken and is calculated between two features and based on its value. Correlation is broadly categorized into three categories as follows:

Positive correlation:

If two features are related in such a way that if one of them increases then the other also increases or if one of them decreases then other one will also decreases then this is called as positive correlation.

Negative correlation:

This correlation occurs when one decreases other increases.

No correlation:

This occurs if there is no relationship between the two features (i.e. the features are independent. So take in consideration of those features which are independent to form input vector for the model.

- **Step 3:** In this step we must pass the input vector to the model for training. After training, the model will then divide the data of the customer into homogeneous segments.
- **Step 4:** After the model is trained we should pass it to the test data to check its accuracy and efficiency.
- **Step 5:** After all this the predictive Model will now predict segments of future customer data.

Phase 2: Extracting the associative buying pattern of segmented Customers

Once the customers are segmented according to phase 1 we find associative buying pattern using association mining and appropriate technique from the particular segments to profit the business.

III. METHODOLOGY

The main purpose of the project is to analyse the customer based on various factors:-

- Age
- Gender
- Purchase history
- Wish lists
- Frequent visits
- Most searched
- Purchase range
- Purchase frequency

Based on the data collected, we categories the customers and understand their requirements. This helps us in satisfying customer expectations. There are mainly two benefits of the project, it helps us understand the customer and suggest them products that might be useful to the customer and eventually increasing the business. These suggestions play a big part in marketing the product.

The other benefit is, it helps us generate offers/coupon codes to the products that the customer desire to buy. The better we understand the customer the better we are able to generate offers. The ultimate goal is marketing, the better we understand our customers the better services we will be able to provide them. Further taking consideration to their search history we can analyze the customer more deeply and provide them with better search results.

There are four common types of segmentation:

1) Demographics
2) Geographic
3) Behavioral
4) Psychographics

Demographic: This is the most common type of segmenting the customer by considering parameter such as age, gender, occupation, income, marital status to create customer segments. It's much easier to get data for demographic segmentation than the other types of customer segmentation .We simply ask the customer to fill forms to get these data.

Geographic: This involves segmenting the customer based on their country, climate, state, region .In this segmentation we must keep the local culture and weather a factor for providing coupons and deals to the customer via messages. By doing so, it provides a greater value to consumer and encourages buying them.

Behavioral: This segmentation is implemented by the basis of how the customer interacts with your brands. This is a the category of segmentation which studies the behavioral traits of consumer, their likes /dislikes, they're response to the product they bought, the kind of brand they choose and promote.

This segmentation is mainly to understand the desire of the customer and to service those products to them.

Psychographic: Just understand ding the customer needs and interests wouldn't suffice, we must know the stage of the buying process they are in, i.e. their current social status. This segment is based on dividing the customers based on psychological factors, which includes behaviors, personalities, lifestyles and beliefs. This segmentation is mainly used to predict how the customer will respond to focused marketing campaigns.

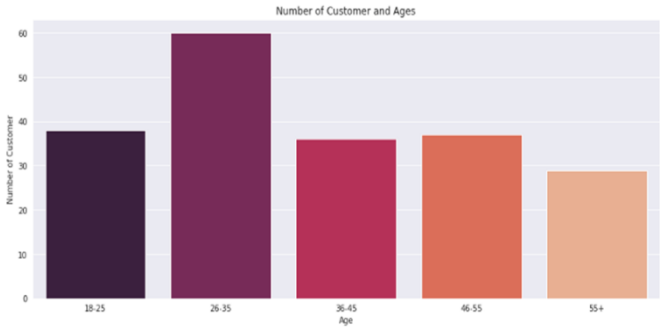

Fig. 2: Result Analysis

The K mean algorithm is discussed as follows:

K-Means is probably the most well-known clustering algorithm in machine learning. It is being taught in a lot of introductory data science and machine learning classes.

1. To begin, we first select a number of groups to use and randomly initialize their respective center

points. To identify the number of groups to use, it's good to take a quick analysis at the data and try to identify any distinct groupings.

2. The data point is classified by computing the distance between that point and each group center, and then classifying the point whose center is closest to the group center to be in the group.
3. Based on these classified points, we recomputed the group center by taking the mean of all the vectors in the classified group.
4. Keep repeating these steps for a set number of iterations or until the group centres don't change much between iterations. There is also an option to opt to initialize the group centres in a random way in a few times, and then select the run that has given best results which is expected.

On the other hand, K-Means has some disadvantages. Firstly, you have to select how many groups there are. This isn't ideal with a clustering algorithm we'd want it to figure those out for our analysis because the point of it is to gain some insight from the data collected. It may also starts with random choice of cluster centers and because of that it may give different clustering results on different runs of the algorithm. Due to this the results aren't repeatable.

IV. CONCLUSION

Customer segmentation is a way to understand and analyze the customer in order to provide customer satisfactory services, but the ultimate goal is satisfying customer so the company's business is maximum. We provide a variety of services like, product suggestion, unique offers to each individual, coupon codes, etc. We categorize each customer with a unique clustering algorithm. The data Is collected based on each individual search history, their wish list, etc. The ultimate goal is to help a company make a successful business

V. REFERENCES

[1]. Al-Qaed F, Sutcliffe A. Adaptive Decision Support System (ADSS) for B2C E-Commerce. 2006 ICEC Eighth IntConf Electron CommerProc NEW E-COMMERCE InnovConquCurr BARRIERS, Obs LIMITATIONS TO Conduct Success Bus INTERNET. 2006:492-503.

[2]. Mobasher B, Cooley R, Srivastava J. Automatic Personalization Based on Web Usage Mining. Commun ACM. 2000;43(8).

[3]. Cherna Y, Tzenga G. Measuring Consumer Loyalty of B2C e-Retailing Service by Fuzzy Integral: a FANP-Based Synthetic Model. In: International Conference on Fuzzy Theory and Its Applications iFUZZY.; 2012:48-56.

[4]. Magento. An Introduction to Customer Segmentation. 2014. info2.magento.com/.../

[5]. An_Introduction_to_Customer_Segmentation...

[6]. Baer D. CSI : Customer Segmentation Intelligence for Increasing Profits. SAS Glob Forum. 2012:1-13.

[7]. Kishana R. Kashwan, Member, IACSIT, C.M.Velu, "Customer Segmentation using clustering and Data Mining Techniques", International Journal of Computer Theory and Engineering, Vol.5, No.6, December 2013.

[8]. R.Kaur ,K.Kaur, "Data Mining on Customer Segmentation:A Review", International Journal of Advanced Research in Computer Science,Volume No-5,2017

[9]. Mark K.Y.Mak,George T.S.Ho,S.L.Ting,"A Financial Data Mining Model for extracting Customer Behaviour", INTECH open access publisher, 23 July 2011

[10]. LuoYe, CaiQiuru, XiHaixu,LiuYijun and Zhu Ghuangping , "Customer Segmentation for Telecom with the k-means Clustering Method", Information Technology Journal 12(3):409-413,2013 .

[11]. I. S. Dhillon and D. M. Modha, "Concept decompositions for large sparse text data using clustering," Machine Learning, vol. 42, issue 1, pp. 143-175, 2001.