

Implementation of Customer Segmentation using Integrated Approach

Balmeat Kaur, Pankaj Kumar Sharma

Abstract: Customer segmentation and pattern extraction is one of the key aspects of business decision support system. In order to grow the business intelligently in competitive market, identification of potential customer should be done timely. This paper proposes an integrated novel approach for determining target customers using predictive model and also discover their associative buying patterns using Apriori algorithm. After identification of targeted customers and their associative buying pattern, the business managers take the strategic profitable decisions accordingly.

Keywords: CRM, Clustering, Customer Segmentation, Associative Mining

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. The most valuable asset of any business is customer. In this emerging market it is very difficult to maintain its customer base. To overcome this difficulty every business has to focus on customer segmentation. Customer segmentation means categorizing the customers into same group. It acts as a base for Customer Relationship Management (CRM) for businesses to firstly identify their target customers and work on them individually. There are various existing predictive models which provide information on customer segmentation and which help them to segregate the customers.

For a successful business, apart from retaining and adding new customers, making more profit from each customer is key task. Different variety of models exist which help the business managers to implement the strategic policy according to individual customer taste but each one of it are having their own limitations. To excel the business further this paper is proposing a new integrated approach of customer segmentation combined with association mining of different segmented customer which provides more profit to the business [3].

The paper is organized as following. Section 2 focuses on proposed methodology, section 3 describes the approaches of proposed methodology and section 4 concludes the study of this paper.

II. PROPOSED METHODOLOGY

Identifying right customer and providing right service at right time and treating different types of customers differently is the key to success in business.

So, a predictive model will be used to segregate customers into different groups based on their transactional data. Once the customers are segregated then their associative buying pattern are identified to enhance the profit for the organization future coming customer.

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

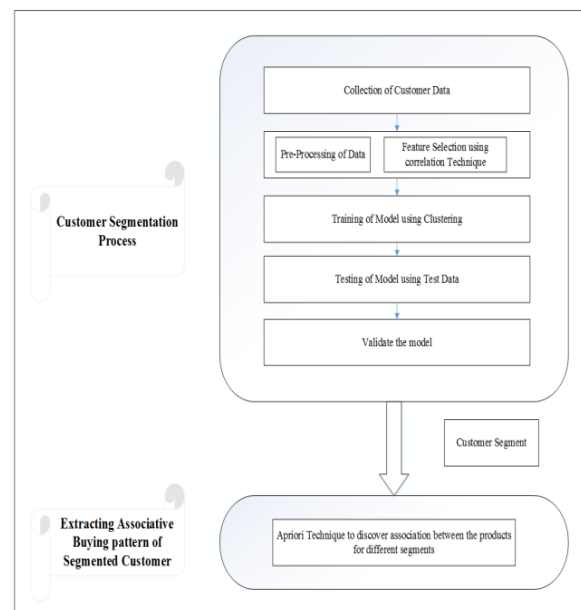


Fig. 1 Integrated Approach of finding customer segments along with their hidden buying patterns

The whole process is to be carried out in two phases

Phase 1: Customer Segmentation

Phase 2: Extracting associative buying pattern of segmented customer

Phase 1: Customer Segmentation

Step 1: Collecting Customer Data (Transactional data):

This step involves the collection of transactional customer data comprises of their static (Eg: Age, Gender etc.) and dynamic data (Eg: Purchase frequency etc.) [1] from shopping vendors.

Step 2: Preprocessing of Data: Pre processing of the data is one of the important step for the accuracy of predictive model. In this step, the collected data will be cleaned and relevant features will be extracted. Feature selection is a data reduction technique which is responsible for extracting relevant features required for input vector of predictive model. This acts as pre-processing steps for creating subset of original features by excluding those features which are redundant.

Revised Manuscript Received on December 22, 2018.

Balmeat Kaur, Assistant Professor, ABES Engineering College
Pankaj Kumar Sharma, Professor, ABES Engineering College

This paper proposes correlation technique for extracting relevant features. Correlation measures the relationship between two features. Basically it simply filters those features which are not redundant to form subset of original features. To measure the association between features correlation coefficient 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 manner that if one increases other also increases or if one decreases then other also decreases then this is called as positive correlation.

Negative correlation: This correlation occurs when one decreases other increases.

No correlation: It occurs when there is no relationship between two features (i.e. the features are independent). So consider those features which are independent to form input vector for model.

Step 3: Pass the input vector to the model for training. After training the model will

Divide the customer data into homogeneous segments.

Step 4: Once the model is trained we pass the test data to check its accuracy and efficiency.

Step 5: Now the predictive Model will predict segments of future customer data.

Phase 2: Extracting associative buying pattern of segmented Customers

Once the customers are segmented in phase 1 we find hidden associative buying pattern using association mining Apriori technique from the particular segments to excel the profit of organization which is discussed in next section.

III. APPROACH FOR PROPOSED METHODOLOGY

Research Method for customer Segmentation

In order to identify the target customers Clustering technique can be used for cluster analysis. Clustering is defined as to group data in clusters/segments so that data within segment are similar while data across the segments are dissimilar. Various techniques can be used for clustering like k means, hierarchical, grid based model based technique. In this paper we proposed to use K-means technique for customer segmentation due its following advantages:

- This technique suits for the data with numeric features and often terminates at local optimum [6].
- It is highly scalable and efficient for large data sets [6].
- It is fast in modelling and its result is more understandable [6].

K means Algorithms is discussed as follows

The k-means clustering algorithm divides the n records into k segments of records called clusters where $k \leq n$, so as to minimize the distances between records within a particular cluster.

Step 1: Choose K points at random as segment centres (centroids).

Step 2: Assign each record to its closest segment centre using certain distance measure (usually Euclidean or Manhattan).

Step 3: Calculate the centroid of each segment, use it as the new segment centre (one measure of centroid is mean).

Step 4: Go back to Step 2, stop when segment centres do not change any more.

Research Method for extracting associative Mining

In order to grow the business more in competitive market, businesses target on cross selling and up selling of the products as well. Cross selling means selling of multiple products offered by a single product to a new customer or old customer and up selling to sell upgraded version of current product. Thus identification of multiple related products for a single product is a key task. This task can be done using the approach Apriori Technique:

Apriori is an iterative level wise search algorithm which is used for finding frequent itemsets using candidate generation[9].

It finds the frequent item-set based on following property:

- If an item-set is frequent, then all of its subsets must also be frequent.
- Support of an item-set never exceeds the support of its subsets .This is called as the anti-monotone property of support

Apriori Algorithm works as follows

Let the set of frequent item-sets of size k be F_k and their candidates be C_k .

In Apriori Algorithm firstly the whole database is scanned and frequent itemsets of size 1 are searched and the count for each item is accumulated and those items are collected which satisfy the minimum support requirement. Then iterates following three steps to extracts all the frequent item-sets.

1. Generate C_{k+1} , candidates of frequent item-sets of size k +1, from the frequent item-sets of size k.
2. Scan the database and calculate the support of each candidate of frequent item-sets.
3. Add those item-sets that satisfy the minimum support requirement to F_{k+1} [9].

IV. RESULT ANALYSIS

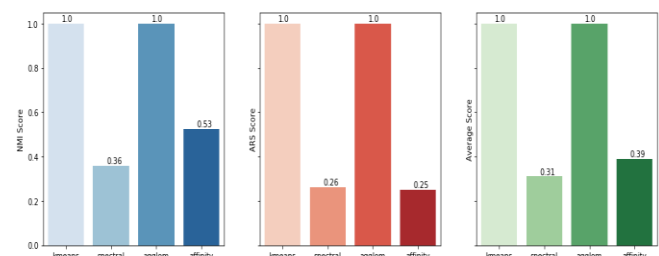


Fig. 2 Analysis of clustering algorithms based on ARS Score and Average Score

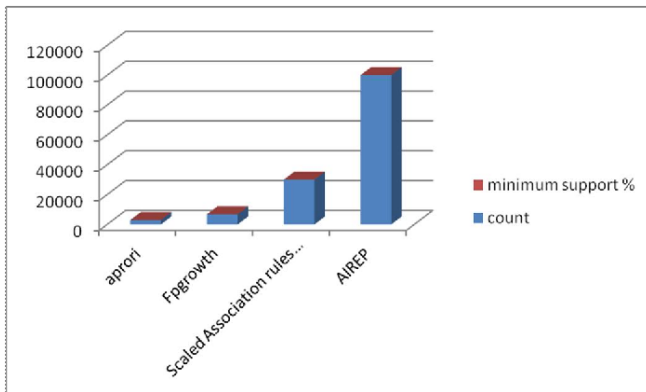


Fig. 3 Analysis of Apriori and other association mining algorithm

V. CONCLUSION

In competitive market of e-commerce, the problem of identifying potential customer is gaining more and more attention. To address this problem timely, this paper proposes a study on integrated novel approach based on clustering using K-means and associative mining using Apriori technique. After identification of targeted customers and their associative buying pattern, the business managers take the strategic profitable decisions accordingly. This integrated model could be directly brought into implementation for providing better profitable margins from sales.

REFERENCES

1. XIANG-BIN YAN, YI-JUN LI, Customer Segmentation based on Neural Network with Clustering Technique, 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 2006 (pp265-268).
2. Ling Luo, Bin Li et. al. "Tracking the Evolution of Customer Purchase Behaviour Segmentation via a Fragmentation-Coagulation Process", Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17).
3. 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.
4. Juni Nurma San, Lukito Nugroho, Ridi Ferdiana, P.InsapSantosa, "A Review on Customer Segmentation Technique on E-Commerce", Advance Science Letters, Vol.4, 400-407,2011.
5. 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.
6. 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 .
7. Yi Zuo,A B M Shawkat Ali , KatsutoshiYada, "Customer purchasing behaviour using statistical learning theory",18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems-KES2014.
8. Haiying Ma, "A study on Customer Segmentation for E-Commerce using the Generalized Association Rules and Decision Tree", American Journal of Industrial and Business Management , 2015,5,813-818
9. Jyoti, Savita Bisnoi " A Predictive Analytics of Cluster using Associative Techniques Tool", IJRDET, Vo -5,Issue 6, June 2016
10. R.Kaur ,K.Kaur, "Data Mining on Customer Segmentation:A Review", International Journal of Advanced Research in Computer Science, Volume No-5,2017