International Journal of Information Technology and Management

Vol. X, Issue No. XVI, August-2016, ISSN 2249-4510

CONSUMER TRENDS AND BEHAVIOR MAPPING WITH DATA MINING FOR ORGANISED RETAIL IN KOLHAPUR DISTRICT

AN INTERNATIONALLY INDEXED PEER REVIEWED & REFEREED JOURNAL

www.ignited.in

International Journal of Information Technology and Management
Vol. X, Issue No. XVI, August-2016, ISSN 2249-4510

Consumer Trends and Behavior Mapping with Data Mining for Organised Retail in Kolhapur District

Mr. Chandrashekhar Shankar Shinde[1]* Dr. Ali Akbar Bagwan[2]

[1]Computer Science & Engineering Department, Kalinga University, Naya Raipur

[2]Computer Science & Engineering Department, Kalinga University, Naya Raipur

Abstract – In this paper we discuss the result of consumer trends and behavior mapping using data mining for retail in Kolhapur. The purpose of this paper is to study, consumer trends and behavior mapping with data mining for organized retail in Kolhapur district as well as implement and analyze various Data-mining algorithms, tools and techniques and then do an analysis of the raw data to obtain a meaningful interpretation. Some of the data mining algorithms we have used are a vector quantization based clustering algorithm, and then an 'Apriori' based Association rule mining algorithm. The first one is aimed at a meaningful segregation of the various customers based on their RFMT values, while the latter algorithm tries to find out relationships and patterns among the purchases made by the customer, over several transactions. Here we also proposed E-cleaner algorithm that use for increase the accuracy of RFMT model.

Keywords—Data mining, RFMT, Aprior, E-Cleanneri.

----------X----------

1. INTRODUCTION

Retail is the offer of products to end clients, not for resale, but rather for utilize and utilization by the buyer. Retail includes the offer of stock from a solitary purpose of procurement straightforwardly to a client who means to utilize that item. The single purpose of procurement could be a physical retail location, an Internet shopping site, an inventory, or even a cell phone. The retail exchange is toward the end of the Manufacturers offer huge amounts of items to retailers, and retailers end to offer those same amounts of items to shoppers. Retailers are the last connection in the inventory network amongst makers and shoppers. Retailing is vital on the grounds that it permits makers to concentrate on delivering products without being diverted with the tremendous measure of exertion that it brings to interface with the end-client clients who need to buy that merchandise.

Retailers make the buy of products simple for the customer. That is the reason retail locations have salesmen, why Internet shopping sites have client benefit moment visit popup, and why indexes have portrayals, photographs and without toll telephone numbers.

Data Mining is the way toward investigating information from alternate points of view and condensing it into helpful data. Data Mining is otherwise called Knowledge Discovery Data and Knowledge Mining (Kendall, 2007). Data Mining is additionally characterized as mining of learning from immense measure of information. With the utilization of Data Mining, nature and conduct of any sort of information can be anticipated. In retail industry, Data mining gathers use amount of data from on deals/buy, client buying history, utilization and administrations. Data mining in retail industry helps in distinguishing clients purchasing examples and patterns that prompt to be enhanced nature of client administrations and fulfillment. Step by step the new advances are assuming its own particular profitable position which meets the cost of substitute way to deal with advertising systems.

There are much more data mining models such as clustering segmentation classification estimation and predictive modeling affinity sets or association protocol illustration and the visualization is the well serious modeling are proposed by number of authors. Similarly there much also application techniques involving association protocol serious pattern grouping analysis classification analysis and probability heuristic study. Thus through analysis knowledge of the customers extracted through information mining can be combined with retailing and then supplied to retailers. The studying the

retailers customer functional shopping information is substantially more subjective in light of the fact that such information more often than not is both numerical and clear cut and the needs of each element exhibiting the information are not characterized, in this examination our thought is to contemplate the information utilizing various unmistakable perception method to imagine client useful information and to speak to the representation comes about for the choices on the procedure of the publicizing effort to a non-specialized individual for the most part the media organizer information on the online shopper shopping capacities is multidimensional and include the two kind number and all out different multidimensional information perception systems are connected in demonstrating the internet shopping data or information and their importance is consider. Now a day almost every work is completed on internet so to make our nation as digital we have large challenge to take concentrate of customer on online shopping and also try to improve the profit on companies which supplied this type of facility.

The remainder of the paper is organized as follows. In section II, we are discussing the related works. In section III discus Proposed work. In section IV, the practical results and analysis is presented. Finally in section V the conclusion is presented.

2. RELATED WORK

This section presents the review of previous methods precisely.

Dr. Sumankumar et. Al (2011)

In [1] has created e-following in India: its issues, openings and compelling procedures for development and advancement, this paper endeavors to give a reasonable picture about the e-following in India and its different issues, openings. It additionally endeavors to draw a viable retailing system in India in view of the point by point study of e-following organizations. The Indian online retail is a rich section holding up to be abused. Web is a powerful medium that can serve as an extraordinary stage for the development of retail brands in India. Web based retailing or e-following is depicted as exchanges that are led through intuitive online PC frameworks, which connect shoppers with venders electronically, where the purchaser and shipper are not at the same physical area. In a short space of time, web retailing or e-following has immovably settled itself as a suitable other option to store based shopping.

Piyushet al (2003)

In [2] has created improvement of present day retailing in India: it's effects on appropriation and acquisition arranges and changing utilization design in this paper the Indian retail industry is assessed to be $470 billion. The composed or cutting edge retailing with 6% share remains at $26 billion. The retailers in various arrangements, including on the web, has influenced an irreversible change in purchaser purchasing propensities and organizations identified with serving these retailers. Be that as it may, not at all like numerous different nations, India has its own particular special character of size, land spread, social assorted qualities, and variety of configurations being presented at brisk progression.

Sunitha et al (2013)

In [3] Retailers additionally confront the test of creating e-following – the mantra of advanced retailer's prosperity this paper web is a strong medium that can serve as a novel stage for the development of retail brands in India. The Indian retail market is seeing unrest. The present electronic models for retailing are a piece of an embryonic stage going before a time of quick change, test, and opportunity in Indian retail advertise. it is not only the metros that are fuelling the online scene in India the request supply hole in level 2-3 urban communities where there is brand mindfulness yet no accessibility of items and administrations is likewise adding to development, Internet retailing entryways, for example, eBay. In, snapdeal.com, and naaptol.com are enlisting anyplace somewhere around 40 and 60% of their deals from provincial territories separated from the level ii and iii urban areas.

Sivathanu (2013)

In [4] has created challenges influencing the sorted out retail segment in this paper the Indian retail industry has been tossed open to outside majors and is stuffed with players who endeavor to offer awesome items and esteem for-cash to Indian purchasers .this exploration paper gives the present status of the retail part in India and the point by point data about the difficulties confronted by the retail segment in India. The paper incorporates the chances of retail locations, retail organize in India and the late patterns in retail. This paper studies the supervisors to comprehend the difficulties affecting the retail part. It finishes up with giving the definite open doors accessible to the retail area.

Sashi Bhushan (2007)

In [5] in hotel industry customer satisfaction is generally an endless supply of administration. Consumer loyalty in the hotel business has been critical for a long time. Superb administration and upgrading consumer loyalty are broadly perceived as critical components prompting the achievement of organizations in the hotel, providing food and tourism ventures. An administration approach concentrated on consumer loyalty can enhance client steadfastness, along these lines expanding the positive picture of the touristic goal. Thus, investigating the significance for clients of hotel properties in lodging choice is crucial. Visitor fulfillment converts into the thought of regardless of

whether clients will come back to a hotel or guidance it to different voyagers is critical to the achievement of the cordiality business.

Jyoti (2013)

In [6] has gotten prospect of e-retailing in India, this paper customer's recognition in regards to shopping has been changed with the presentation of web media. The electronic retailing (e-following, e-retailing, web retailing and so on.) is the model of offering of retail merchandise utilizing electronic media, specifically, the web. E-retailing is a subset of ecommerce (electronic business). E-retailing represents around 10% of the general development of ecommerce market. The development in the e-retailing business sector is driven by the need to spare time by urban India. It is evaluated that 2.5 billion web clients, access to web has assumed a huge part in developing the business markets. The web gives retailers an instrument for: expanding target markets, upgrading purchaser connections, amplifying product offerings, enhancing cost effectiveness, enhancing buyer interchanges, and conveying modified offers.

Meenakshi (2013)

In [7] concentrate on development of retail market in India with exceptional reference to widening of shopping center culture in level ii city this paper now the shopper needs to shop at a place where he can get nourishment, excitement and shopping, all under one rooftop. This has given Indian sorted out retail advertised a noteworthy support. Shopping centers are the hot shopping goals in new way of life. Because of quick retail blast, assortment of shopping centers has developed. With this first-class shopping center culture hitting the colossal Indian working class, the times of unadulterated shopping delight is by all accounts reducing.

Rajora et al (**2013)**

In [8] provincial retailing in India – an evolving worldview, this paper in the realm of merciless rivalry, each advertiser is looking better arrangements and items and administrations to serve the end clients, around 70 % populace is lives in rustic or semi urban range and as of late, country market of India have gained noteworthiness, as the general development of the Indian economy has brought about the considerable increment in the obtaining force of the country groups. Truth is told it has been evaluated that the provincial markets are developing at five times the rate of urban markets, in this way offering sufficient chances to advertisers. In this paper analyst will talk about the points of view in the rustic retail& challenges in the provincial showcasing and a few proposals about how country advertising should be possible all the more adequately.

Bortiz et al. Verhoef (2003)

In [9] analyze the study on data mining for CRM showed that data mining method used to explicit untapped useful knowledge from a huge amount of customer data. The most important goals of Data mining are predicting and building describing knowledge. CRM generate intercommunication of customers with the organization by using information technology (IT). Moreover specifying customer demand or interest better and treating them accordingly can enhance their lifetimes. Customer clustering is the sets of customers into distinct sets based on their common attributes it is the main part of CRM.

Dunham (2003)

In [10] award data mining used in build six kinds of models went for illuminating business multifaceted nature order relapse time arrangement bunching affiliation investigation and arrangement disclosure. The initial two arrangements and relapse are utilized to make forecasts. Affiliation and grouping disclosure are utilized to exhibit conduct. Bunching can be utilized for either gauging depiction. One of the essential sorts of prescient displaying assignments is order in characterization based administered learning information are mapped into classes or gatherings which are re-characterized before the information are inspected.

3. PROPOSED METHODOLOGY

In this section, the proposed method architecture and algorithms are presented.

The purpose of this methodology is to improve customer trends and behavior mapping with data mining for organized retail in Kolhapur district as well as implement and analyze different Data-mining algorithms, tools and methods and then do an examination of the sample/raw data to acquire a meaningful interpretation. Few of the data mining algorithms we have utilized are a vector quantization based different clustering algorithm, and afterward an 'Apriori' based Association rule mining algorithm. The first one is aimed at based on their RFMT values a meaningful separation of the various customers is done, while the latter algorithm tries to calculate out patterns and relationships among the purchases made by the customer, over several transactions.

Data Organization Layer:- In this layer, we collect data like that is history of user transaction , gather data are maintain in a customer's transaction wise so that the data must be represent as a time serie**s.** Our system accepts user transaction list dataset which contain name of item purchase by customer in each transaction. Using that we cleaner method to clean

data set in which we remove all unwanted data as well as calculate support of each item set and crate hyper clique pattern and finally remove those item which not follow that pattern like that we clean all dataset. This layer mainly focuses on dataset, collection of dataset and how to improve quality of dataset. We select historical data which contains all transaction history of user from all big market in Kolhapur district.

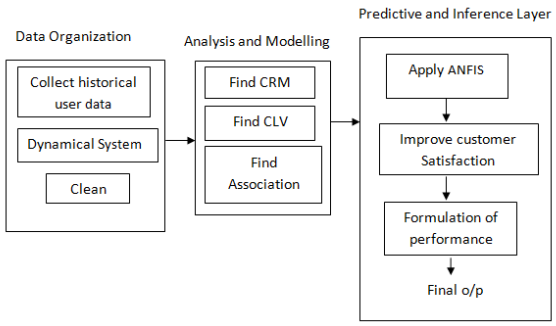

Figure 1 System Architecture

Analysis and Modeling: - In this phase we concentrate on find customer behaviour and relationship. So first we calculate CRM. Customer relationship management (CRM) is a theoretical concept that use to technologies, practices and strategies that organized utilized to handle and study data throughout and customer behaviour the customer lifecycle. The main goal of CRM is increasing business association with customers, assisting in customer driving and retention sales growth. To compile data on customers across different channels, CRM systems are create or points of contact with the customer and the organization which could include the organization's website, direct mail, telephone, live chat, social media and marketing materials. These systems can also giving customer-facing staff detail information on customers' personal information, buying preferences purchase history and concerns. Then we go through CLV (Customer Lifetime Value). The customer lifetime value equation essentially views a customer as an income stream. So instead of considering the customer's purchases as single transactions, the marketing focus becomes creating ongoing series of profitable transactions. To measure performance of CLV we use RFTM model. All the result of RFMT with comparison of RFM model respect to CLV is disused chapter 4. These ongoing transactions are created through customer relationship management practices and strategies – with the success of CRM activities being measured by improvements in the firm's customer lifetime value. Finally find association between item set to calculate relation between products so it's helpful to predict product. To calculate the associations we use Apriori algorithm through association mining rules are create. For finding interesting relations between customers in large databases Association rule learning is use which is a rule-based machine learning method. To detect the most important relationships association rules are created by studying data for frequent if or then patterns and using the criteria confidence and support. How regularly the items appear in the database is indicated by support. The number of times the if or then statements have been found to be true is shown by confidence.

Predictive and Inference Layer: **-** Finally in predictive and inference layer probabilities are utilized to get specific desired information. In this we user ANFIS (Adaptive Neuro Fuzzy Inference System) to predict selling product which will increase profit of business so we have to predict those product in that customers are surely interested to buy. ANFIS is fusion of Fuzzy Inference Systems (FIS) and Artificial Neural Network (ANN). Then we focus on improve customer satisfaction by implement CRM. We predict top most products to user as well as this layer focus on performance like time and accuracy of prediction which improve the performance as compare to existing system. So if performance is increase then there is huge chance to increase online shopping trends and trend increase propositionally profit also increase. A predictive mining approach imagines sales for another space relies upon the present information. The enormous issue show in the investigation of offers in view of the conditions among the items and the client division, which expands the matter of the retail locations. To recognize client's conduct another procedure is proposed. The proposed procedure is relies upon the blends of information mining methodologies, for example, grouping and affiliation govern mining. It focuses on the finding of standards and focuses on showcasing the items relies upon the populace. Affiliation rules made for an area at a state of offer can't be more compelling in a portion of the area since the entire and complex conduct of customer and their procedures in choosing things are unique.

In below algorithm, we eliminate all unwanted fields which not required detecting attacks.

Data cleaning is nothing but the data scrubbing or cleansing, deals with deleting and finding errors and inconsistency from information in order to increase the data quality. We proposed the algorithm namely Extended hyper clique based data cleaner (E-Cleaner). Total working of E-Cleaner can be summarized as follows:

E-Cleaner Algorithm:

Input: Transaction set T

Result: Set of points which is not noise NP, Set of noise points NO.

for i = 1 to n do

if T[i].contain(Special_characters)then

T[i].remove(Special_characters)

```
end

HCL ← HypercliqueMiner(T) //HCL: the hyperclique set;

T[1...n].covered ← false;

num hc ← size(HCL);

for i = 1 to num hc do

for j = 1 to n do

if ((!T[j].covered)&&contains(T[j],HCL[i])) then

T[j].covered ← true;

end for j

end for i

NO ← {};

NP ← {};

for i = 1 to ntrans do

if T[i].covered then

NP ← NP ∪ T[i];

end

else

 NO ← NO ∪ T[i];

end

end
```

Firstly we do pre-processing on each transaction in which we remove html characters, stop words, expressions, URLs and punctuation and then find hyper clique patterns as per given h-confidence threshold and support threshold and after that eliminate such objects that are not available in present hyper clique pattern. For any data set the hyper clique patterns depends upon the value of both support and h-confidence thresholds. When we have to possible, we set the default values for support threshold to be near to zero and the h-confidence threshold use to manage the all number of objects that are consider as noise. Therefore, we use a low support threshold that is higher sufficient to minimize the total number of hyper clique patterns to a manageable level. This algorithm is utilized to minimize process time of proposed system so it will help full to increase output of system.

4. RESULTS AND DISCUSSION

Below figure shows accuracy of proposed model it shows higher because of the proposed H-cleaner algorithm.

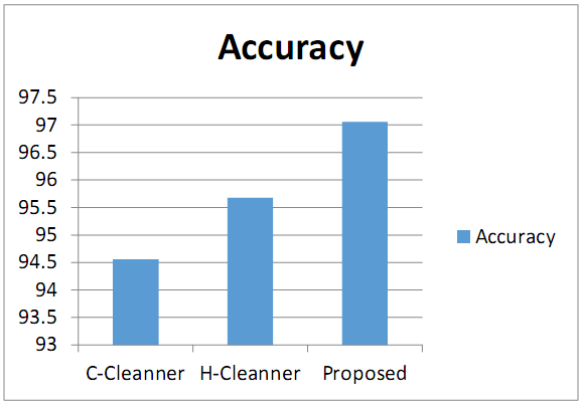

Figure 2: Performance Analysis of Accuracy

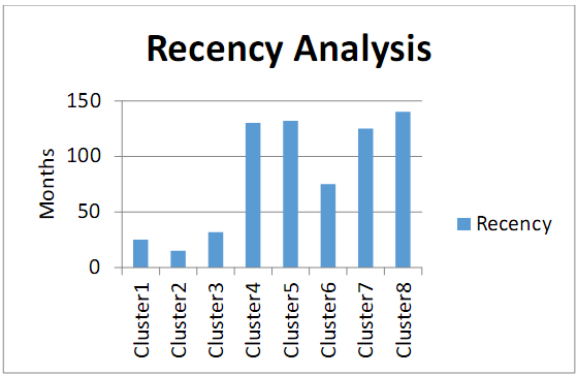

Figure 3: Regency analysis of cluster

The regency analysis of each cluster is shown in figure 2. The frequency analysis of each cluster is shown in figure 3. The monetary analysis of each cluster is shown in figure 4. The term analysis of each cluster is shown in figure 5.

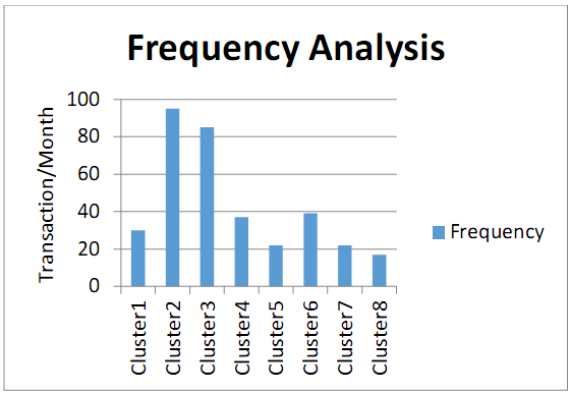

FIGURE 4: FREQUENCY ANALYSIS OF CLUSTER

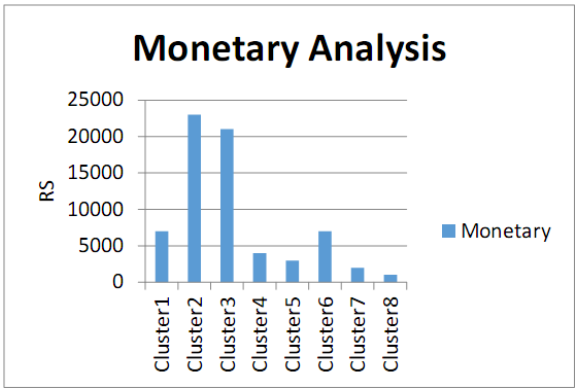

FIGURE 5: MONETARY ANALYSIS OF CLUSTER

5. CONCLUSION

In this thesis we have identified Customer trends and behaviour with reference to Customer Lifetime Value, CRM and Customer Satisfaction using data mining techniques. The proposed system try to increase the customers satisfaction, identifies customer interest and increase profit of retailer at the earliest possible point. To examine and remove unwanted data E-Cleaner is used. It is used to remove extra space or special symbol as well as transaction which not useful for further process. RFMT model is used to evaluate customer's loyalty and is used to assess customer's lifetime value. To discover customer knowledge ANFIS model has been proposed. The proposed system is designed to improve customer satisfaction. Clustering is use to classify Customer data. Customer's behaviour can be identifying using association rules so using that rules we can recommend proper product to user so there is more possibility to buy product. The main goal of this thesis is to highlight how Marketing professionals and retailers can use the data mining tools to move closer to their customers and add value to their products. It suggest the potential for applying data mining tools and technology to find out the best fit methods for pattern discovery and matching.

REFERENCES

Arora Jyoti (2013). "Prospect of e-retailing in India." IOSR journal of computer engineering, volume 10, Issue no.3, pp. 11-15.

Bansal Mr. Puneet, Ms. Veerpaul Kaur Maan, and Mr. Mandeep Rajora (2013). "Rural Retailing in India–A Changing Paradigm." International Journal of Advanced Research in Computer Science and Software Engineering, volume 3, Issue no.11, 2013.

Bortiz et. al. (2003). Verhoef, CRM (Customer Relationship Management) data mining. GTL information Technology and Organization.

Chandra, P. Sathish, and G. Sunitha (2010). "E-TAILING-THE MANTRA OF MODERN RETAILER's SUCCESS." Researchers World, Volume 3, Issue no. 2. pp. 42, 2010.

Choudhary, Meenakshi (2013). "Growth of Retail Market in India with Special Reference to Broadening of Mall Culture in Small Townships." International Journal of Innovative Research and Development, volume 2, Issue no. 2, pp. 547-566.

Dawn, Dr. Suman Kumar, and Uttiya Kar (2011). "E-Tailing in India: Its Issues, Opportunities and Effective Strategies for Growth and Development." International Journal of Multidisciplinary Research, volume 1, Issue no. 3.

Dunham, Weiss and Indurkhya (2003). The characteristics of human genes: analysis of human chromosome, Published online in wiley inter-science 4: pp. 635-646.

Kendall, Stephanie D. (2007). "Customer Service from the Customer's Perspective". In Fogli, Lawrence. Customer Service Delivery: Research and Best Practices. J-B SIOP

Sinha Piyush Kumar (2003). "Shopping Orientation in the Evolving Indian Market", Vikalpa, Vol. 28, Issue no. 2, pp. 13 – 22, April - June, 2003.

Sudame P., Sivathanu B. R. (2013). "Challenges affecting the organized retail sector", AJ of Econ and Mgmt., Volume 2. 2013.

Corresponding Author

Mr. Chandrashekhar Shankar Shinde*

Computer Science & Engineering Department, Kalinga University, Naya Raipur

E-Mail – csshinde7769@gmail.com