

Data Mining Techniques to Improve Customer Relationships Management

Siddhesh Chavan¹, Aditya Jadhav², Prashik Suryagandh³, Prof. Nidhi Sharma⁴

^{1,2,3} Student, Department of Computer Science Engineering Bharati Vidyapeeth College of Engineering, Navi Mumbai

⁴Department of Computer Science Engineering Bharati Vidyapeeth College of Engineering, Navi Mumbai

Abstract - Advancements in today's technologies have made huge relationship markets in recent years. Today's competitive world requires to manage customer relationship. Technologies such as data mining, data warehousing and campaign management software have made customer relationship a new area to deal with. Data mining is the process in which variety of techniques and models are used to find the relationships and the patterns related in different types of data, so that we can make accurate decisions using this patterns. The whole system purely depends on the decisions so the decision making process has much more importance in the field of data mining. The involvement of information technology revolution, and particularly in web area the opportunity to make better CRM increases. The system replaces traditional available offline systems. A business might use a data and create its own information to work. This information helps to construct a CRM model to deal with new emerging technologies. Here using previous available data is processed and new future decisions are generated.

Key Words: Customer Relationship Management, Data items, K-means.

1. INTRODUCTION

1.1 Customer Relationship Management

CRM is a generation for managing all enterprise relationships and interactions with customers and ability clients. The purpose is easy: enhance business relationships. A CRM machine helps organizations stay linked with clients, streamline tactics, and improve profitability. When human beings communicate about CRM, they're typically referring to a CRM device, a device that enables with touch management, sales control, productivity, and greater. A CRM solution facilitates you attention for your employer's relationships with individual people — which include clients, service users, colleagues, or suppliers — all through your lifecycle with them, including finding new clients, triumphing their enterprise, and providing help and additional offerings all through the relationship. According to Microsoft, "CRM is a consumer-focused enterprise method designed to optimize sales, profitability and consumer loyalty through implementing a CRM strategy, an enterprise can improve

the commercial enterprise techniques and generation answers around selling, marketing, and servicing functions across all purchaser contact-factors (for example: net, email, cell phone, fax, in-person)". The overall goal of CRM applications is to attract, retain and control a firm's profitable ("right") clients. Digital aid for such activities is furnished through the firm's commercial enterprise intelligence machine. The essence of the data generation revolution and, particularly, the arena of extensive web is the opportunity to construct higher relationships with customers than has been formerly feasible in the offline international. until these days maximum CRM software has centered on simplifying the organization and control of customer records. Such software program, referred to as operational CRM, has targeted on creating a customer database that prevents a steady image of purchaser's courting with the agency, and offering the statistics in unique applications but, the sheer volumes of customer records and an increasing number of complex interactions with customers have propelled records mining to the vanguard of making your customer courting profitable. CRM refers to strategies and technologies hired by using enterprises to address their relationships with customers. records collected on cutting-edge clients (and prospective clients) is evaluated and used to the ship. computerized patron dating control strategies are frequently used to create automated customized advertising based on patron facts from the device. CRM is the enterprise-level tactic, targeting making and keeping relationships with customers. Numerous enterprise CRM software solutions have distinctive method to CRM, then again, CRM isn't always a technology, but fairly a holistic technique to an corporation's point of view, having the significance on the client resolutely. CRM direct an company's standpoint at all levels, together with guidelines and approaches, customer service, personnel education, systems and information control, advertising. CRM solutions have integrated non-stop advertising, income and consumer services. A CRM machine must classify elements essential to clients, propose a purchaser-oriented method, receive purchaser based measures, provide higher customer support, extend lower back-to-lower back procedures for customers/clients, song all elements of sales, take care of consumer grievances, generate a holistic evaluation of customer service and so forth. A CRM device regularly follows a customer shopping

for behavior and necessities, in addition to their pursuits. This statistics can be used to purpose customers selectively. Additionally, the product a consumer has bought may be attempted in the course of the product life cycles, letting customers to acquire records concerning a product to purpose clients with facts on different products once a product starts to be phased out.

1.2 Data Mining

Data mining, the extraction of concealed analytical data from huge databases, is an effective and new technology with awesome potential to assist organizations by concentrating on the most essential data in the data warehouse and data marts. Data mining tools and techniques anticipate future patterns and practices permitting organizations to make positive, proactive and learning-drive choices. The computerized, prospective examinations or analyzes proposed by data mining move past the investigations of past events offered retrospective tools. Data mining tools can respond to business questions that were excessively tedious, making it impossible to resolve. Most enterprises collect and refine huge amount of data. Data mining strategies can be executed quickly an existing hardware and software platforms upgrades the benefits of existing data assets, data resources and can be integrated with new products and frameworks as they are brought online. Data mining tools can analyze the huge amount of data in the databases by using high performance client/server or parallel processing computers. It delivers answers to inquiries such as, "Which clients are most likely to respond to my next promotional mailing, and why?"

2. SYSTEM APPROACH

2.1 Integrating CRM and Data Mining technologies

2.1.1 Creation of customer database

A first step is necessary to complete CRM solution which is construction of customer/End user database. This will be the basis of any CRM activity. In the business which are Web-based it is relatively straightforward activity as the customer transactions and the customer details are collected as a main part of communication with customers. The companies which are working in this particular field they are³ collecting data about customers. This task consists of getting historical customer data from their internal sources. The question is : What to be extracted from the existing data? Basically database consists of huge data which contains following information:

1. Transaction/Activity: Keeps tracks of the customer activity.
2. Customer Details: Now-a-days customers are available from many categories, so to it is important to have most of the information about customer/End user.
3. Customer connections: Customer contacts and the referencing to the other customers give a boost to the business as it does not require company interference.

2.1.2 Data Analysis

Traditionally, the database of customers are analyzed with the purpose to define and segregate the customers segments. Various multivariate statistical methods such as cluster and discriminant analysis have been used to make a community of a community of customers with similar behavioral patterns and descriptive data which then utilized to develop different product offering or a direct marketing campaign. Mostly direct marketers utilize such techniques for many past years. Their aim is to target the most preferable prospects for catalogue mailing to adapt the catalogue suitable for different customers.

2.1.3 Selection of Customers

"The process of customer elimination is also termed as customer selection" which has huge importance in CRM strategies. As CRM is a purely customer oriented business strategy and hence we pays attention to improve responsiveness to the changes in customer requirements. Serving all customers may not be important as many of the customers are occasional users, Hence the organizations working over this strategy will have to consider set of customer to reduce time and also workload. So that business could grow efficiently. The result from the analysis phase could be of different types. Customer contributing the following categories are normally selected first for the further process (e.g., highest purchasing rates, greater brand loyalty etc). If individual customer-based profitability is available through LCV process or any similar analysis the determination of customer selection becomes easier process. Marketing manager is responsible to make various kind of analysis to focus a particular customers. The goal is to make use of customer profitability, which will provide long term assurance of the profit to organization. These customers also provides information to the manager to fire the customer which is expensive to serve relative to that of the revenues produced by them.

2.1.4 Focusing the Customers

Traditional approaches such of mass marketing like television, radio, or print media and advertising are useful in creating awareness and achieving other communication objectives, but in today's world they are considered as slow communication channel Thus, they are poorly suited for the CRM system because of their unbiased nature. The ordinary approaches for targeting selected customers consist of direct marketing methods like telemarketing, mail, direct messages. When the nature of product is suitable direct sales strategy is also preferred. Because of the internet the new strategy is adopted by the

organizations to work in a 1-to-1 marketing with customers. Personalized E-mails are the most preferred way in direct marketing.

2.1.5 Relationship Packages

While customer contact through direct Email services is a useful element of CRM. Relationships aren't built and sustained with direct electronic mails themselves but instead through the sorts of packages that are available for which Email can be a communication mechanism. the overall intention of relationship packages is to supply a higher stage of purchaser satisfaction. Managers nowadays sense that clients in shape realizations and expectations of product overall performance, and that it's miles important for them to deliver such performance at higher and higher ranges as expectancies increase because of excessive or heightened opposition, and changing customer desires. similarly, studies has shown that there may be a sturdy, fine relationship among customer delight and earnings. Managers need to continuously degree pleasure stages and increase programs that assist to supply performance past focused consumer expectancies

2.1.6 Confidentiality

Those CRM framework characterized depends upon An database about client records and assessment of the information for All the more capable and centered on promoting Furthermore advertising interchanges Also relationship-building exercises. There might be a evident tradeoff "around those ability of associations should finer convey custom planned items Furthermore administrations and the amount of facts necessary with tolerance this transport. Uncommonly with the Ubiquity of the internet, A large number clients What's more backing aggregations need aid concerned around the measure about particular information held done databases what's more entryway it may be getting utilized.

2.1.7 Metrics

The multiplied interest paid to CRM means that the conventional metrics used by managers to degree the success of their products and services inside the market need to be up to date. Economic and market-based totally indicators like profitability, marketplace share, and income margins have been and could stay essential. but, in a CRM world, extended emphasis is being located on developing measures which are client-centric and supply the supervisor a higher concept of the way the CRM rules and packages are working.

2.2 Customer Retention Program

2.2.1 Customer Service

As because clients have more picks nowadays and the focused clients are most treasured to the organisation, customer support should receive a high priority inside the

agency. packages designed to beautify customer support are usually of two types. Reactive service is where the purchaser has a trouble (product failure, question about a invoice, product return) and contacts the organization to clear up it. most agencies these days have hooked up infrastructures to address reactive service conditions. Proactive service is a distinct count; that is a situation where the manager has decided now not to await customers to touch the firm however to alternatively be competitive in organising a communicate with clients previous to complaining or other conduct sparking a reactive solution.

2.2.2 Loyalty/Frequency programs

Loyalty applications (also known as frequency packages) provide rewards to customers for repeat shopping. a number of internet-primarily based companies providing incentives for repeat visits to net websites. even though these have not been wildly a success, it's far clear that the rate orientation of many net customers creates the need for applications that could generate loyal conduct.

2.2.3 Customization

The thought about impostor customization may be setting off secret word 1-to-1 promoting Also showcasing as a result it infers those prologue about results Furthermore administrations for mamoncillo or lady clients, never again truly talking to them. Customization is termed "versioning." it is, from claiming route, simpler should do this to administrations Also immaterial holding detail products over for stock however the illustrations over indicate that Significantly Producers camwood make get of the developed Realities accessible starting with clients should tailor stock that in any event convey those presence for continuously custom outlined regardless of they might make sincerely variants ahead a as a relatable point build.

2.2.4 Reward programs

Various organizations used to present rewards to its clients on successive buying of their items. praise may be a present object or coins back or like bargain coupon, which may be carried out on buying the next object from the identical organisation.

2.2.5 Community Building

One of the principal uses of the internet for both online and offline groups is to build a community of clients for exchanging product-associated statistics and to create relationships between the clients and the organisation or brand. these networks and relationships are known as communities. The aim is to take a prospective dating with a product and turn it into something extra private. on this way, the supervisor can build an surroundings that makes it more difficult for the customer to depart the "family" of different people who also purchase from the organization.

3. SYSTEM IMPLEMENTATION

3.1 K-means clustering Algorithm

K-means is signal processing originated method of vector quantization, which is a famous method for cluster analysis in data mining. This algorithm aims to partition n observations into k clusters which every observation belongs to the cluster with the nearest mean. This process makes the partitioning of the data space into Voronoi cells. There are heuristic algorithms that are commonly used and combined quickly to find a local optimum. K-means clustering finds clusters of comparable spatial extent.

3.2 Description:-

Given: Set of observations $\rightarrow (X_1, X_2, \dots, X_n)$.

Aim: To partition the no of observations into k clusters (k \rightarrow no.of clusters). [Note:k<=n]

The technique also minimizes the within cluster sum of square also known as variance.

Output: sets S = (S₁, S₂, ..., S_n).

$$\mu_i \parallel^2 = \arg \min \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 = \arg \min \sum_{i=1}^k |S_i| \text{Var } S_i$$

$\mu_i \rightarrow$ mean of point S.

$$\arg \min \sum_{i=1}^k \frac{1}{2|S_i|} \sum_{x, y \in S_i} \|x - y\|^2$$

The Equivalence can be deduced from identity:

$$\sum_{x \in S_i} \sum_{x \in S_i} \|x - \mu_i\|^2 = \sum_{x \neq y \in S_i} (x - \mu_i)(\mu_i - y)$$

3.2 Algorithm:-

Input: no. of clusters \rightarrow k

set of dimensions \rightarrow D

Output: set of k clusters

Procedure:

Randomly choose k dimensions from D as initial cluster centers;

Repeat:

1. (Re)assign each object to the cluster to which the object is the most similar, based on the mean value of the objects in the cluster;

2. Update the cluster means, i.e. calculate the mean value of the objects of each cluster Until no change;

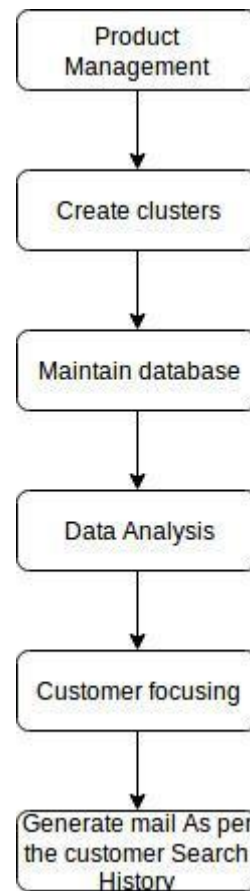


Fig 3.1: System Flow

4 TECHNICAL APPROACH

At the time of insertion of the products in the shopping cart we assign a particular code for a category/domain. Then those categories also contain classification according to the brand. The actual product id consists of following different fields (category_id, brand_id, product_id). Now, from the customer browsing history by taking customer action in consideration we make a customer database, which maintains the data and the records about the customer interest. After mining database we get brief idea about the customer requirements and the interest of customer. After the mining process automatic email generation module takes the control and sends the mail to the customer. In this way the customer relationship is maintained using technology. It makes the process very simple and efficient. Also reduces paperwork and helps to maintain the environment. The system reduces the human intervention, Hence there won't be any chance of human error. The system provides the exact results, Hence it helps to boosts the profit of the entrepreneurs.

5. REQUIREMENTS

5.1 Hardware Requirement

Hardware	Specification
Processor	Intel core 2 duo or above
RAM	2GB

5.2 Software Requirement

Software	Specification
Operating System	Microsoft Windows 7/8/10 or Unix
Platform	Xampp/lampp server

6. CONCLUSION

Nowadays, the e-commerce and e-marketing is very important concept as the whole market has been controlled by using internet technologies. To maintain a healthy relationship with customers and he end users the e-commerce comprises a lot. Such types of tools and technologies are been used by e-commerce companies. This leads to increase the profit and business also.

REFERENCES

- [1] J.Han and M. Kamber, Data Mining: Concepts and Techniques (2nd ed.). Morgan Kaufmann, 2006.
- [2] Ngai, E. W. T., Xiu, L., & Chau, D. C. K., "Application of data mining techniques in customer relationship management: A literature review and classification," Expert Systems with Applications, 2009, 2, 2592–2602.
- [3] Parvatiyar, A., & Sheth, J. N. "Customer relationship management: Emerging practice, process, and discipline", Journal of Economic & Social Research, 2001,3, 1–34.
- [4] Shu-hsien Liao, Yin-ju Chen, Hsin-hua Hsieh, "Mining customer knowledge for direct selling and marketing", Expert Systems with Applications, 2011, 38, 6059–6069.
- [5] Rong-Shiunn Wu, Po-Hsuan Chou., "Customer segmentation of
- [6] multiple category data in e-commerce using a soft-clustering approach" Electronic Commerce Research and Applications, Volume 10, Issue 3, May–June 2011, Pages 331-341.
- [7] Kim, Y. H., & Moon, B. R., "Multicampaign assignment problem. IEEE Transactions on Knowledge and Data Engineering", 2006,18, 405–414.
- [8] Hyea Kyeong Kim, Jae Kyeong Kim, Qiu Yi Chen "A product network analysis for extending the market basket analysis", Expert Systems with Applications, Volume 39, Issue 8, 15 June 2012, Pages 7403-7410
- [9] García, V., Sánchez, J. S., & Mollineda, R. A., "On the effectiveness of preprocessing methods when dealing with different levels of class imbalance", Knowledge-Based Systems, 2012,1, 13–21.
- [10] Quinlan, J. R. C4.5: Programs for machine learning, Morgan Kaufmann, 1993.
- [11] Russell, Stuart; Norvig, Peter Artificial Intelligence: A Modern Approach (2nd ed.).Prentice Hall. 2003.
- [12] Vladimir N. Vapnik, The Statistical Learning Theory. Springer, 1998.
- [12] D.W.Aha, D.Kibler, K.Albert, "Instance-based learning algorithms", Machine Learning,1991, 6, 37-6.