International Journal of Advance Research in Science and Engineering
Vol. No.5, Issue No. 05 , May 2016
www.ijarse.com
IJARSE
ISSN 2319 - 8354

OPTIMUM SOLUTION FOR CUSTOMER RELATIONSHIP MANAGEMENT BASED ON TWO PHASE CLUSTERING ESTIMATION

M.Rajakumar[1] , V.Jayaraj[2]

[1]*Research Scholar, Dept, of Computer Science, Periyar University, Salem, (India)*

[2]*Associate Professor, Bharathidasan University, Thiruchupalli, (India)*

ABSTRACT

This paper deals with providing an optimal solution that optimizes the process of Customer Relationship Management. The process is initially performed by using the Clustering technique and when a new customer arrives, he is classified into any one of the existing Clusters which automatically gives us the properties for the customer. Here we perform a two phase Clustering which not only clusters on the basis of a single attribute, but performs a secondary clustering operation for better and accurate results. Further, all the Clustered information are indexed in such a way that Clustering becomes faster.

Keywords: Customer Relationship Management, Indexing, Customer Value, Clustering, Classification, Two Phase Clustering.

I. INTRODUCTION

Customer Relationship management is the strongest and the most efficient approach in maintaining and creating relationships with customers. Customer relationship management is not only pure business but also ideate strong personal bonding within people. Development of this type of bonding drives the business to new levels of success.

Once this personal and emotional linkage is built, it is very easy for any organization to identify the actual needs of customer and help them to serve them in a better way. It is a belief that more the sophisticated strategies involved in implementing the customer relationship management, the more strong and fruitful is the business.

CRM has been initially performed manually by employees of the organization. Due to the increase in online transactions, it becomes better and efficient to automate this process [8], [9].

A proper mathematical model cannot be proposed for this strategy, since this process becomes unpredictable, as it involves human being as an important element.

When this type of scenario is encountered, a continuous learning algorithm is a key to success [4]. Our paper proposes such an algorithm integrating the elements of Data Mining.

II. USE OF DATA MINING TECHNIQUES IN CRM

Data mining, which is also called KDD (Knowledge Discovery in Database), is the process of abstracting unaware, potential and useful information and knowledge from plentiful, incomplete, noisy, fuzzy and stochastic actual data. Simply speaking, it is a process to pick up the information and knowledge which can not be discovered directly but with potential value from a mass of data. Here, we use the process of Clustering, Classification and Neural Networks for the mining data.

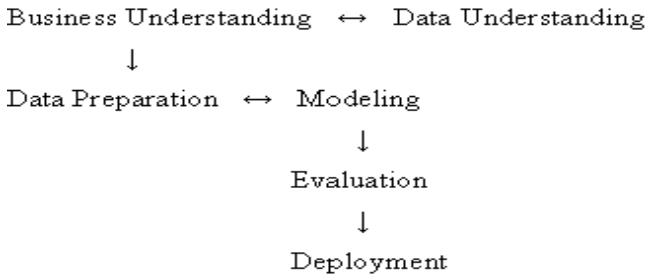

Fig. 1: Data Mining Process

2.1 Neural Networks

An artificial neural network (ANN), often just called a "neural network" (NN), is a mathematical model or computational model based on biological neural networks, in other words, is an emulation of biological neural system. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation.

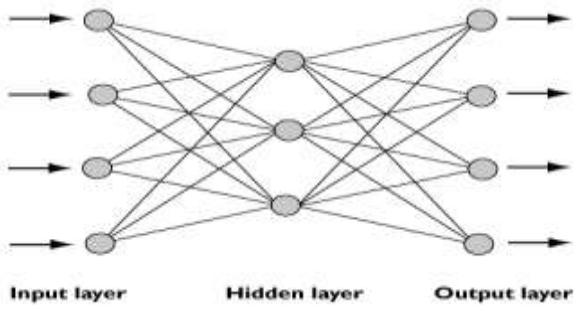

Fig. 2: Structure of an Artificial Neural Network

In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase.

An ANN is typically defined by three types of parameters:

1. The interconnection pattern between different layers of neurons
2. The learning process for updating the weights of the interconnections
3. The activation function that converts a neuron's weighted input to its output activation.

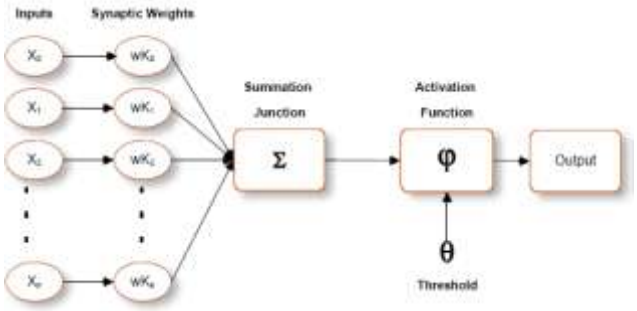

Fig. 3: Mathematical Model of a Neural Network

2.2 Clustering

Data clustering is a method in which we make cluster of objects that are somehow similar in characteristics. The criterion for checking the similarity is implementation dependent. Data Clustering is a technique in which, the information that is logically similar is physically stored together.

2.3 Classification

Data classification is the categorization of data for its most effective and efficient use. In a basic approach to storing computer data, data can be classified according to its critical value or how often it needs to be accessed, with the most critical or often-used data stored on the fastest media while other data can be stored on slower (and less expensive) media.

By applying these techniques in CRM, the process of customer analysis becomes more accurate.[2],[3],[7]

III. THE CRM PROCESS

The CRM process is carried out in two different phases. The first phase deals with processing of the data for the first time, and further processing are carried out according to the rules mentioned in the second phase.

The first phase involves gathering the required data from the user and performing the operation of Clustering on the obtained data. This results in producing clusters of similar types of data, as described in [1]. We further refine this

process by performing clustering operation on the already clustered data, which provides us a further degree of refinement [10]. After the successful completion of this process, the obtained clusters are analyzed for common information. This information is recorded along with the common properties used for Clustering process. This forms the basis for Indexing.

When a new customer arrives, he/she is classified into any one of the existing clusters. The property of the customer is said to correspond to the properties of the cluster in which they are classified into.

The second phase deals with providing the user a faster performance during the clustering process. After a certain period of time, the data present in the data store increases and hence a need for re-clustering arises. This process is usually tedious, since the entire first phase is carried out with a larger amount of data. But in the proposed paper, the data present in the clusters are indexed, since obtaining of properties from an indexed format is usually faster than the normal database, we can save a large amount of time. Only new data present in the system have to be analyzed and stored in the index. This saves a considerable amount of time and processing power.

Both, the clustering and the classification phases are performed using the artificial neural networks.

3.1 Phase I

The historical purchase data [6] that is available for the organization under study is taken as the training data for the neural networks. [4] Each property of the customer is fed into a perceptron. The perceptrons are already provided with weight-age values that correspond to the property represented by it.

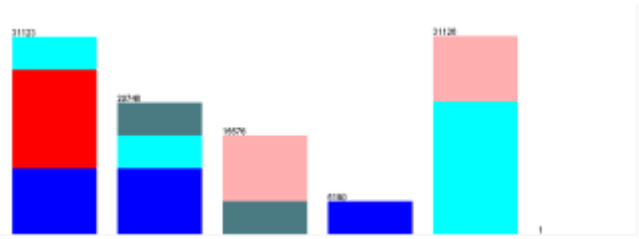

Fig. 4: Histogram representing the salary values of customers(Output From Weka 3.0)

The second hidden layer consists of one perceptron that inputs all the values provided by the first layer. All these properties are processed and a final value called the Customer Value [5] is obtained from the system.

A similar process is carried out for all the products that are being sold by the organization and a product value is obtained for each product.

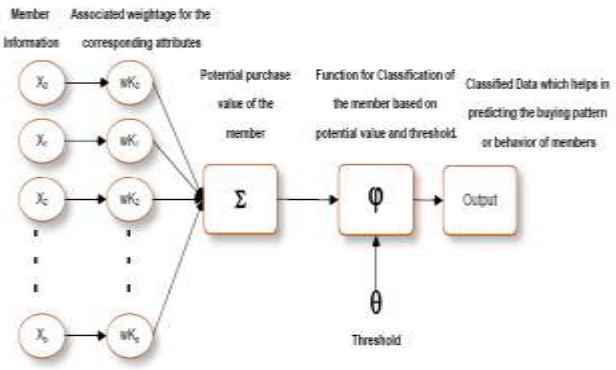

Fig. 5: User Profile Evaluation (Process of Clustering)

The minimum and maximum customer values are obtained from the result and a median value is calculated. This value acts as the base for Clustering. A boundary is defined by the user, with the obtained median value taken as the centroid. All the customer values that come under this boundary are added to the cluster.

A second phase Clustering is carried out on each of these obtained clusters as described in [10]. These clusters will correspond to a new property provided by the customer. This provides us with an additional division in each cluster or a second degree of refinement. Each cluster is examined and similar properties are extracted out of each cluster and are considered as the properties belonging to the cluster. This property directly corresponds to the purchase capacity of the customer, or we can also say that these properties reflect the type of products that a customer most likely purchases. A similar process is carried out for the products cluster and product properties are obtained. All these data obtained are recorded in an indexed table. This table contains the customer value, the property values that were used in the clustering process and the cluster under which the current customer is present.

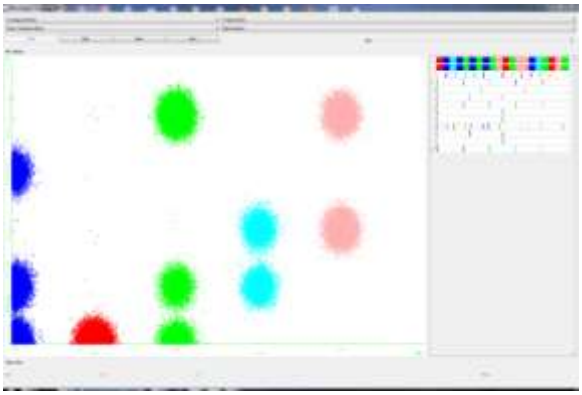

Fig. 6: Clusters corresponding to salary and purchase category (Output from Weka 3.0)

When a new customer data is provided to the system, the customer's properties are passed to the perceptrons and they are passed to the first level hidden layer. This layer contains the final weight-age values that have been tuned in for the particular system. These weight-age values are processes with the corresponding properties and the results are passed to the second level hidden layer. The second level perceptrons processes this cumulative information and the customer's value is obtained. The obtained customer value is compared with the existing clusters and it is classified onto a cluster that has the current customer value within its boundaries. This process is called Classification. [9]

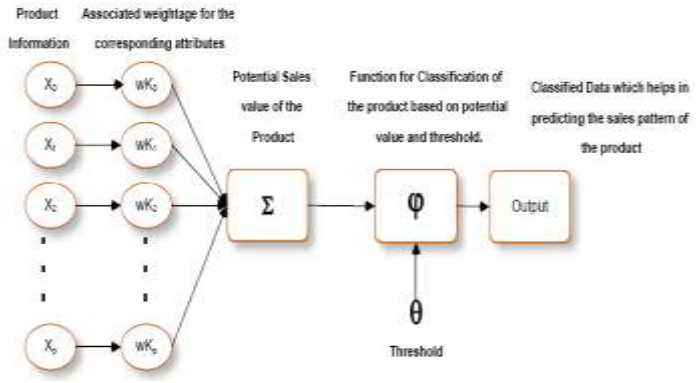

Fig. 7: Process of Classification

The properties of the Cluster, under which the particular customer is classified, are taken as the properties corresponding to the customer. These properties are compared with the properties that have been generated from the product clusters. The two clusters are integrated to provide a solution to the user that contains all the products that the customer has the highest probability of buying.

3.2 Phase II

After a period of time, the database is prone to expand. When this happens, the user is compelled to carry out the clustering process again. This process keeps getting tedious every time this stage is reached, since every time a larger amount of data is to be handled by the user. This can be solved by using indexing. Instead of referring the entire database, the user can refer to the indexed values alone, which is faster. The new values are added to the table every time this process is carried out. Hence, processing is done only for the newly arrived data.

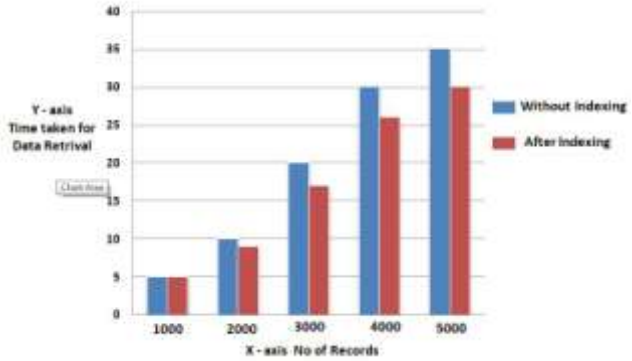

Fig. 8: Comparison Time Chart with and without Indexing

IV. CONCLUSION

This project provides an enhancement for the previous paper [1]. Even though [1] gives us a generic solution, it is still a tedious process every time the processing is carried out.

The running time for phase I is found to increase periodically, hence it cannot provide an optimal solution for extended periods of time. Using the currently proposed method, the running time can be reduced to a large extent. Hence the usability of the system is found to increase.

Since a two-way clustering is performed on the data, the currently proposed project provides a better solution.

REFERENCES

[1] Dr.V.Jayaraj, J.Jegathesh Amalraj, Enhancing Customer Relationship Management Based on Customer and Product Value Estimation - A Neural Networks Approach, International Conference on Networks, Intelligence and Computing Technologies ICNICT'11, (2011).

[2] Wanghualin, Yuzhongdong, The Research of Customer Knowledge Management in CRM, International Conference on Intelligent Computation Technology and Automation (2010).

[3] Gao Jie, Customer Value Based On Data Mining, International Conference on Educational and Network Technology ICENT, (2010).

[4] Smita Nirkhi, Potential use of Artificial Neural Network in Data Mining, 978-1-4244-5586-7/10, IEEE, (2010).

[5] Xu Ling, Song Li, CRM - Customer Value Based On Data Mining, Third International Conference on Knowledge Discovery and Data Mining (2010).

[6] Gu Yu, Li Jiahui, Hierarchy analysis and optimization of customer value based on data mining technology, International Conference on Intelligent Computation Technology and Automation, (2010).

[7] Kun Wu, Feng-ying Liu, Application of Data Mining in Customer Relationship Management, 978-1-4244-5326-9/10, IEEE, (2010).

[8] Gao Hua, Customer Relationship Management Based on Data Mining Technique, 978-1-4244-8694-6/11, IEEE, (2011).

[9] Wanghualin, Data Mining and Its Applications in CRM, Second International Conference on Computer Research and Development, (2010).

[10] Namvar, Gholamian, KhakAbi, A Two Phase Clustering Method for Intelligent Customer Segmentation, 978-1-4244-5984-1, IEEE, (2011).

.