Classification and Prediction in CRM Using Back Propagation Multilayer Feedforward Neural Network Approach

Leela Rani .K
Final M.Tech, Dept Of CSE,
Sri Vasavi Engineering College, Pedatadepalli,
Tadepalligudem—534101, India

G Loshma
Associate Professor, Dept of CSE,
Sri Vasavi Engineering College, Pedatadepalli,
Tadepalligudem—534101, India

Abstract— **CRM "is a business strategy that aims to understand, anticipate and manage the needs of an organisation's current and potential customers"**. **Customer Relationship Management provides a customer classification and prediction which is used for the optimization of business process. customers". It is a "comprehensive approach which provides seamless integration of every area of business that touches the customer- namely marketing, sales, customer services and field support through the integration of people, process and technology" This classification and prediction in CRM will help the company to study, analyze and forecast customers pattern of consumption, business transaction and purchasing CRM has become major activity in the enterprise based business organization using the CRM. CRM is an important activity in the enterprise business organization like banking industry, insurance industry, retail industry and manufacture industry. In the system we are using data mining techniques to implement customer classification in CRM as we need to analyze mass volume of data we are implementing an efficient and effective Neural Network based technique. Based on the existing system like Naïve Bayesian System, our proposed system implements Back propagation Neural Network techniques which would generate accurate results with less time complexity.**

Keywords — Classification, Prediction, Data Mining, Naïve Bayesian Algorithm, Back propagation, Neural networks.

I. INTRODUCTION

CRM is through customer segmentation is organized around the business, cultivating a customer-centric operation behavior, the implementation of customer-centric business processes to improve corporate profitability, revenue and customer satisfaction a business strategy.The core part of CRM activities is to understand customer requirements and retain profitable customers. To reach it in a highly competitive market, satisfying customer's needs is the key to business success [1]. Unprecedented growth of competition has raised the importance of retaining current customers. Retaining existing customers is much less expensive and difficult than recruiting new customers in a mature market. So customer retention is a significant stage in Customer Relation Management, which is also the most important growth point of profit [3]. Factors that influence customer satisfaction degree are concerned by all enterprise managers [9]. Marketing literature states that it is more costly to engage a new customer than to retain an existing loyal customer. Churn prediction models are developed by academics and practitioners to effectively manage and control customer churn in order to retain existing customers [4]. So, Customer satisfaction is most important. Data mining (DM) methodology has a tremendous contribution for researchers to extract the hidden knowledge and information which have been inherited in the data used by researchers [7]. Data mining has a tremendous contribution to the extraction of knowledge and information which have been hidden in a large volume of data [8]. The concept of customer satisfaction and loyalty (CS&L) has attracted much attention in recent years. A key motivation for the fast growing emphasis on CS&L can be attributed to the fact that higher customer satisfaction and loyalty can lead to stronger competitive position resulting in larger market share and profitability [6][7].

However, it is a difficult and a complex task to identify the customer's needs such as colors and design of the products. The objective of this paper is to design and implement the expert system in order to assess customer satisfaction and reveal appropriate strategies to improve it. As the customer satisfaction on colors and design can have a complex hidden pattern and, therefore, the approach of the paper should have an ability to perform pattern recognition, classification and forecast which make the artificial neural networks an appropriate technique to be applied in the expert system[9]. The conceptual work of the paper is illustrated in Figure 1, in which the assumption of the customer requirements and expert system are based upon the statement that ''in general, the same customer group will like the same colors''. A vast variety of colors mixing in different products that makes it a difficult and complicated task to identify the customer's needs. The contribution of this paper is in designing the system that is the combination of the expert system and the ANN. The customers can interact with the interface of the expert system to ask and get the advices from the system. Correlation Coefficient can be found. According to that, we can identify the customer's behavior [8][9]. 2

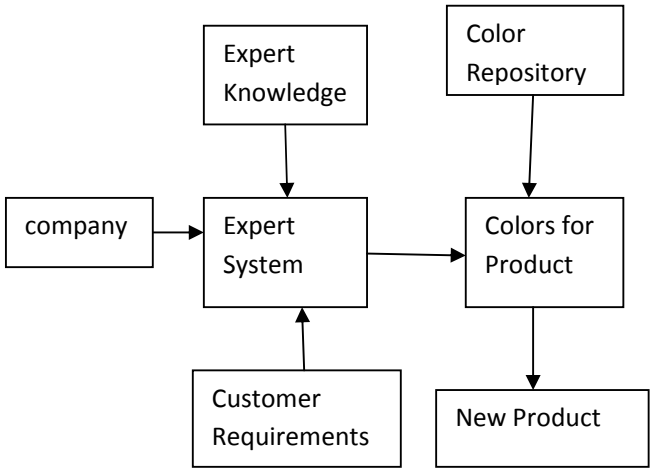


Figure 1.1. The Conceptual Work

This paper is organized as follows: Section II describes related review. Section III describes Architecture of CRM based on Data Mining, Section IV describes proposed work (Methodology) . Section V describes implementation details. Section VI presents experimental results and performance analysis. Section VII presents conclusion and future scope.

II. LITERATURE REVIEW

In the process of economic development, business management concept evolved from product-oriented idea to market-oriented idea, and then to customer-oriented idea. Whether enterprises can obtain, maintain and develop their own clients or not has become the most critical factor because customers are important strategic resources. Customer Relationship Management(CRM) is based on the respect to customers and it can help enterprises understand the entire life cycle of the customers and enable enterprises to provide the customers with more personalized and more efficient services, and then it can enhance customer satisfaction and loyalty, and improve the competitiveness of enterprises ultimately. Customer segmentation, as the core of CRM, is to classify the customers according to the customer's attributes, behavior, needs, preferences, value and other factors in a clear business strategy and specific market, and it can provide appropriate products, services and marketing models to the customers.

A. Data Mining and CRM

The first and simplest analytical step in data mining is to describe the data. For example, you can summarize data's statistical attributes (such as means and standard deviations), visually review data using charts and graphs and look at the distribution of field values in your data. But data description alone cannot provide an action plan. You must build a predictive model based on patterns determined from known results and then test that model on results outside the original sample[1][2].

A good model should never be confused with reality (you know a road map isn't a perfect representation of the actual road), but it can be a useful guide to understanding your business. Data mining can be used for both classification and regression problems. In classification problems you're predicting what category something falls into – for example, whether or not a person is a good credit risk or which of several offers someone is most likely to accept. In regression problems, you're predicting a number, such as the probability that a person will respond to an offer. In CRM, data mining is frequently used to assign a score to a particular customer or prospect indicating the likelihood that the individual behaves the way you want [1][2][3]. For example, a score could measure the propensity to respond to a particular offer or to switch to a competitor's product. It is also frequently used to identify a set of characteristics (called a profile) that segments customers into groups with similar behaviors, such as buying a particular product. A special type of classification can recommend items based on similar interests held by groups of customers. This is sometimes called collaborative filtering.

B. Expert System

The expert system's role is in the preparation to capture the knowledge of the experts and the data from the customer's requirements. The system has the capability to compile the collected data and form the appropriate rules for choosing fragrance notes for the products. In order to identify the hidden pattern of the customer's needs, the artificial neural networks technique has been applied to classify the fragrance notes based upon a list of selected information [3]. The expert system's role is in the preparation to capture the data from the customer's requirements and predict appropriate perfume.

For this end, factors of perfume costumers' decision were recognized using Fuzzy Delphi method and a back propagation neural network classification model was developed and trained with 2303 data of customers [4]. The proposed business intelligent system for demand forecasting proves to give more accurate prediction for future demands compared to the existing models and practices in spare parts inventory management. This helps inventory managers to better manage their supply chain performance by reducing reaching days and service level simultaneously. Reaching day as a measure of inventory level is generally reduced successfully by the retailers at the cost of service level in most of the places [5].

C. Baysian Classification

Bayesian classification is based on Bayes theorem. Studies comparing classification algorithms have found a simple Bayesian classifier known as the naive Bayesian classifier to be comparable in performance with decision and neural network classifiers.

Bayesian classifiers have also exhibited high accuracy and speed when applied to small databases. The naive Bayesian classifier works as follows, as in[7]:

1.Each data sample is represented by an n-dimensional feature vector $\boldsymbol{X = (x1, x2 ,..., xn)}$, depicting *n* measurements made on the sample from *n* attributes, respectively, *A*1, *A*2 ,..., *An* .

2.Suppose that there are *m* classes, $\boldsymbol{C1,C2 ,...,Cm}$ **.** Given an unknown data sample, *X* (having no class label) ,the classifier

will predict that X belongs to the class having the highest posterior probability, conditioned on X .
That is, the naïve Bayesian classifier assigns an unknown sample X to the class Ci if and only if $\boldsymbol{P(Ci\ X) > P(C\ j\ X)}$ for $\boldsymbol{1 \le j \le m, j \ne i}$.

Thus we maximize $P(Ci\ X)$. The class Ci for which $P(Ci\ X)$ is maximized is called the maximum posterior hypothesis. By Bayes theorem,

$$\mathbf{P(Ci/X)=P(X/Ci)P(Ci) / P(X)}$$

3. As $P(X)$ is constant for all classes, only $P(X\ Ci)P(Ci)$ need be maximized. If the class prior probabilities are not known, then it is commonly assumed that the classes are equally likely, that is, $\boldsymbol{P(C1) = P(C2) = ... = P(Cm)}$ and we would therefore maximize $\boldsymbol{P(X\ Ci)}$. Otherwise, we maximize $\boldsymbol{P(X\ Ci)P(Ci)}$.

4. Given data sets with many attributes, it would be extremely computationally expensive to compute $P(X\ Ci)$. In order to reduce computation in evaluating $P(X\ Ci)$, the naive assumption of class conditional independence is made. This presumes that the values of the attributes are conditionally independent of one another, given the class label of the sample, that is, there are no dependence relationships among the attributes. unknown sample using naive Bayesian classification, given the training data as Table 1. The data samples are described by the attributes: sex, age, student and income. The class label attribute, creditcard_proposing has two distinct values(namely,{yes, no}).

III. SYSTEM ARCHIRECTURE

If data warehouse, OLAP and other data analytic tools are directly applied to enterprise's data base, for example sale data base, marketing data base and customer service data base, it will be not only waste time and substance but also not ideality, therefore, it is wise that the data of business data base is lead to data warehouse after was cleaned, extracted, transformed, loaded, the data of data warehouse will offer the best data resource for the application of data mining. The data analysis will become great efficiency based on the data warehouse.The architecture of CRM based on data mining is as follows.

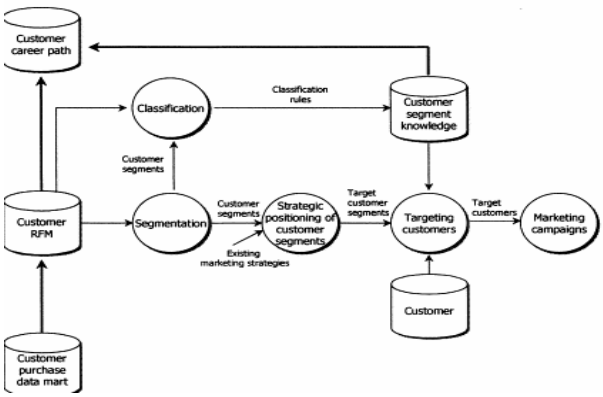


Fig 3.1 The architecture of CRM based on data mining

IV. PROPOSED WORK (BACK PROPAGATION BASED CLASSIFICATION)

A. Classification

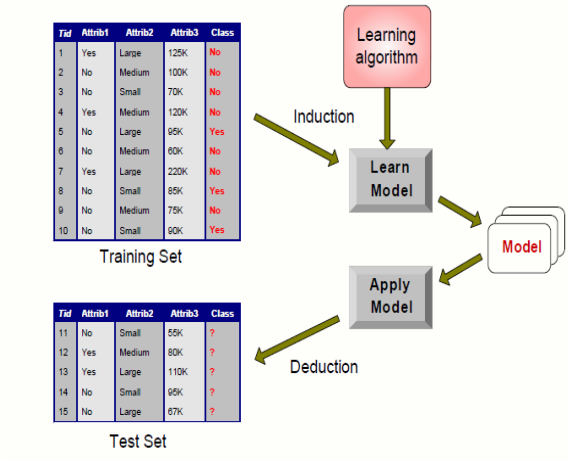
Classification is a data mining (machine learning) technique used to predict group membership for data instances. For example, you may wish to use classification to predict whether the weather on a particular day will be "sunny", "rainy" or "cloudy". Popular classification techniques include decision trees and neural networks. It can be formally defined as Given a collection of records (*training set*)

– Each record contains a set of *attributes*, one of the attributes is the *class*.

Find a *model* for class attribute as a function of the values of other attributes.

Goal: previously unseen records should be assigned a class as accurately as possible.

– A *test set* is used to determine the accuracy of the model. Usually, the given data set is divided into training and test sets, with training set used to build the model and test set used to validate it.



The following diagram illustrates the classification task.

Figure4.1Classification Process

B.Back Propagation

Multilayer perception l & BP (Back-propagation) model Standard multilayer perception (MLP) architecture consists more than 2 layers; A MLP can have any number of layers, units per layer, network inputs, and network outputs such as fig 4.2 models. This network has 3 Layers; first layer is called input layer and last layer is called output layer; in between first and last layers which are called hidden layers. Finally, 4

this network has three network inputs, one network output and hidden layer network.

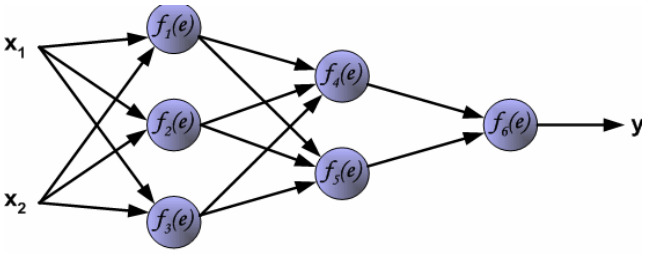
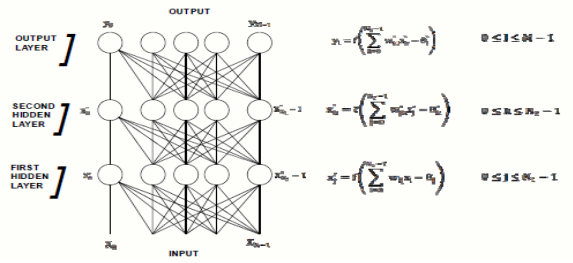


Fig4.2 Standard Multi layer perception architecture



To teach the neural network we need training data set. The training data set consists of input signals (x_1 and x_2) assigned with corresponding target (desired output) z. The network training is an iterative process. In each iteration weights coefficients of nodes are modified using new data from training data set. Modification is calculated using algorithm described below: Each teaching step starts with forcing both input signals from training set. After this stage we can determine output signals values for each neuron in each network layer. Pictures below illustrate how signal is propagating through the network, Symbols $w_{(xm)n}$ represent weights of connections between network input x_m and neuron n in input layer. Symbols y_n represents output signal of neuron n. First, a training input pattern is presented to the network input layer. The network propagates the input pattern from layer to layer until the output pattern is generated by the output layer. If this pattern is different from the desired output, an error is calculated and then propagated backward through the network from the output layer to the input layer. The weights are modified as the error is propagated

Phase 1: Propagation

Each propagation involves the following steps:

1. Forward propagation of a training pattern's input through the neural network in order to generate the propagation's output activations.
2. Backward propagation of the propagation's output activations through the neural network using the training pattern's target in order to generate the deltas of all output and hidden neurons.

Phase 2: Weight update

For each weight-synapse follow the following steps:

1. Multiply its output delta and input activation to get the gradient of the weight.
2. Bring the weight in the opposite direction of the gradient by subtracting a ratio of it from the weight.

This ratio influences the speed and quality of learning; it is called the *learning rate*. The sign of the gradient of a weight indicates where the error is increasing, this is why the weight must be updated in the opposite direction.

Repeat phase 1 and 2 until the performance of the network is satisfactory

Algorithm for Back Propagation

Step1: Initialize weights and offsets
Set all weights and node offsets to small random values.

Step2: Present input and desired outputs
Present a continuous valued input vector X0, X1…..XN-1 and specify the desired output d0,d1,….dM-1. If the net is used as a classifier them all desired outputs are typically set to zero except for that corresponding to the class the input is from. That desired output is 1. The input could be new on each trial or samples from a training set could be presented cyclically until stabilize.

Step 3: Calculate Actual Output
Use the sigmoid non linearity from above and formulas as in fig 3 to calculate output y0,y1….yM-1.

Step 4: Adapt weights
Use a recursive algorithm starting at the output nodes and working back to the first hidden layer. Adjust weights by

$$w_{ij}(t+1) = w_{ij}(t) + n\,\delta_j\,x_i \quad -(1)$$

In this equation Wij (t) is the weight from hidden node I or from an input to node j at time t, ' j w , is either the output of node i or is an input, η is a gain term, and j δ , is an error term for node j , if node j is an output node, then

$$\delta j = y_j(1-y_j)(d_j-y_j) \quad -(2)$$

where d_j s the desired output of node j and y_j *i*s the actual output.

Step 5: Repeat by going to step 2

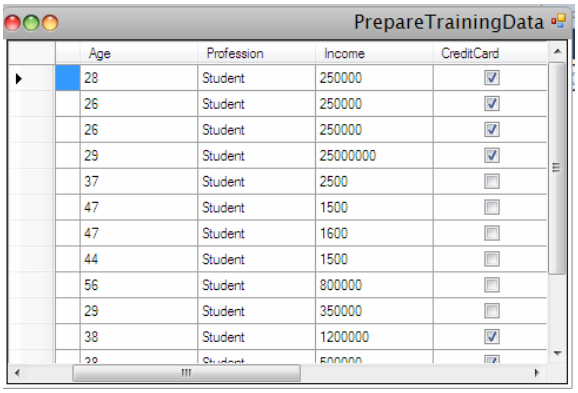
V. IMPLEMENTATION DETAILS

In this section, we analyze the performance of our new approach for classifying the customer credit card data using back propagation approach and predicting the who likely customers to our new business are. This we compare with Bayesian classification approach. The Algorithms were implemented in DOT NET Frame work using C# language. Forms framework is used for designing GUI. We have placed transactional data records in data sets. The SQL Server 2000 data base is used for managing the performance results

.

VI. EXPERIMENTAL RESULTS

In order to evaluate the performance of our proposed algorithm, we have conducted experiments on a PC (CPU: Intel(R) Core2Duo, 3.16GHz) with 4GByte of main memory



running Windows XP. The following shows the results of Back propagation Classification Algorithm for generating customer relationship prediction.

Fig 6.1:Preparing the Training Dataset.

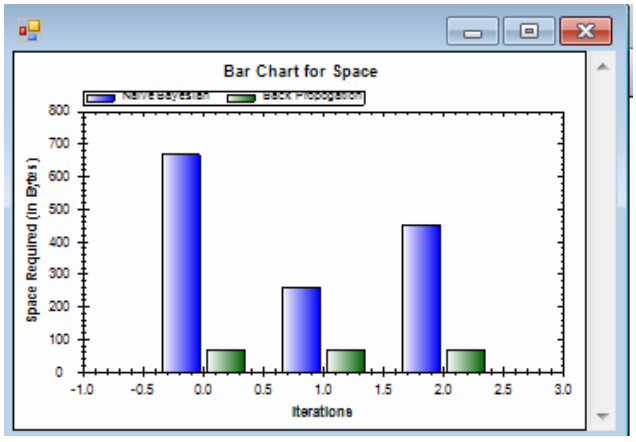


Fig 6.2 Space Complexity for Naïve and Back Propagation Classifier approaches.

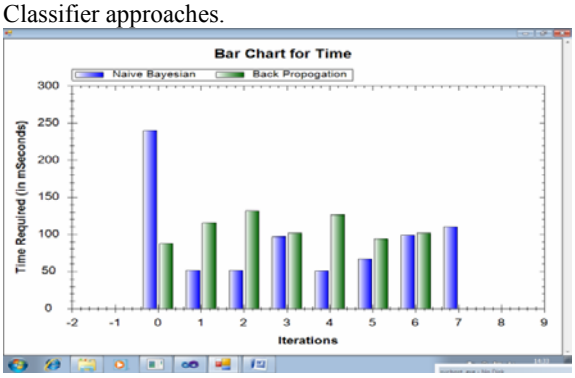


Fig 6.3 Time Complexity for Naïve and Back Propagation Classifier approaches

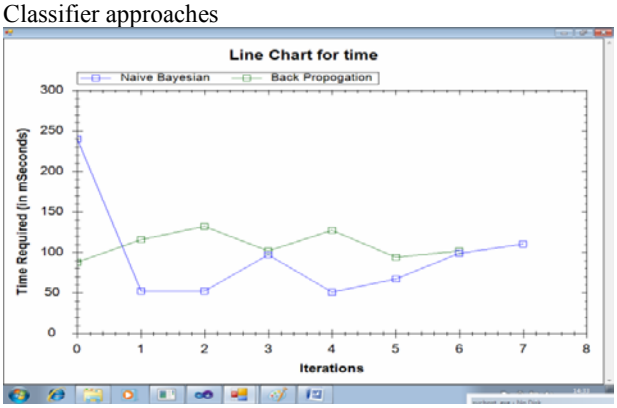


Fig 6.4 Time Complexity for Naïve and Back Propagation Classifier approaches

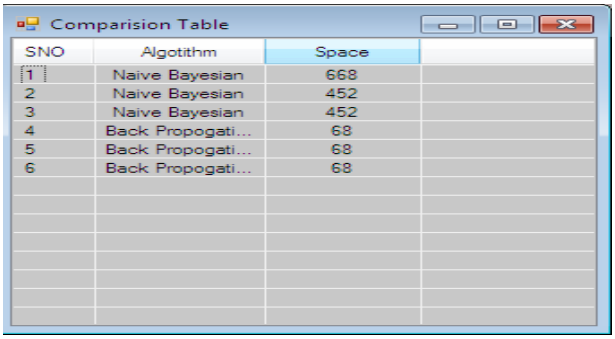


Fig 6.5 Comparision table for Naïve and Back Propagation Classifier approaches

VII. CONCLUSION AND FUTURE WORK

In the dynamic business environment, information systems need to evolve to adapt to new requirements, which may be driven by CRM. The paper has presented a framework of an evolving information system based on knowledge from data mining, and has discussed the framework by focusing on knowledge of classification. Data mining provides the technology to analyze mass volume of data and/or detect hidden patterns in data to convert raw data into valuable information.This paper mainly focused on the research of the customer classification and prediction in Customer Relation Management concerned with data mining based on Back propagation technique. Backpropagation is an iterative process that can often take a great deal of time to complete. When multicore computers are used multithreaded techniques can greatly decrease the amount of time that backpropagation takes to converge. If batching is being used, it is relatively simple to adapt the backpropagation algorithm to operate in a multithreaded manner.

REFERENCES

[1] Gao Hua, 2011 IEEE, "Customer Relationship Management Based on Data Mining Technique",.

[2] Jenkins D(1999). "Customer relationship management and the data warehouse" [J]. Call center Solutions, 892):88-92.

[3] Chung HM and P Gray(1999). "Data mining" [J]. Journal of MIS,16(1):11-13

[4] Balaji Padmanabhan, Alexander Tuzhilin , On the Use of Optimization for Data Mining: Theoretical Interactions and eCRM Opportunities[J], Management Science . Vol. 49, No. 10, October 2003, pp. 13271343.

[5]HOKEY MIN, Developing the Profiles of Supermarket Customers through Data Mining[J], The Service Industries Journal, Vol.26, No.7, October 2006, pp.747–763

[6] Davis B(1999). "Data mining transformed" [J]. Informationweek, 751:86-88.

[7] Kuykendall L(1999)."The data-mining toolbox"[J]. Credit Card Management,12(6):30-40.

[8] Han, J. and Kamber M., (2001) Data Mining: Concepts and Techniques[M], San Francisco, CA: Morgan Kaufmann Publishers.

[9] Kantardzic, M. (2003) Data Mining: Concepts, Models, Methods, and Algorithms[M], Piscataway, NJ: IEEE Press.

[10]. Huang, C.-L., Chen, M.-C., and Wang, C.-J., (2007) Credit scoring with a data mining approach based on support vector machines, *Expert Systems with Applications*, 33,847–856.

[11]. Hsu, C.F., Chang, B., and Huang, H.F. (2007) Applying SVM to build supplier evaluation model – comparing Likert scale and fuzzy scale, in the Proceeding of IEEE International Conference on Industrial Engineering and Engineering Management (IEEE IEEM), 6-10.