International Journal of Applied Engineering Research ISSN 0973-4562 Volume 15, Number 3 (2020) pp. 253-258

Customer Deposit Prediction Using Neural Network Techniques

Ashalata Panigrahi[1] and Manik Chand Patnaik[2]

[1]*Roland Institute of Technology, Berhampur, India*

[2]*Roland Institute of Technology, Berhampur, India*

Abstract

The digital nature of today's banking market, banks collect customers information using different channels such as e-mail, phones such as fixed line or mobile for sharing information about the products or services . Bank offers products such as home loans, personal loans, financial planning, life assurances, retirement plan etc. Multiple decisions are taken by bank in a day for its beneficial impact on customers in a cost effective and timely manner. The decisions are taken with the consideration of customer information such as transaction details, credit card details, risk profiles, saving account, long-term deposit investment. The objective of the work is to predict whether a customer will subscribe a term deposit or not. In this work, we propose neural network based techniques to construct a decision making model using six classifiers, namely, SMO, SVM, RBFN, MP, SOM, and HLVQ. Further, the most relevant attributes in the input data have been selected through a preprocessing stage using three attribute evaluator methods namely, filtered attribute evaluator, one-R attribute evaluator, Relief-F attribute evaluator. Finally, the performance of the model has been evaluated on the bank marketing dataset in terms of accuracy and specificity.

Keywords: Radial Basis Function Network, Self Organizing Map, Sequential Minimal Optimization, Learning Vector Quantization, Multilayer Perceptron.

1. INTRODUCTION

Now-a-days banks are faced many challenges offering innovative products and services to the public such as continually rising marketing cost, decrease response rate. The global banking sector is rapidly changing and developing in which innovations are applied widely. Especially, in recent years, redesigning of working manner and activity structure in the banking sector has become widespread in all around the world. The developments in behavior and preferences of consumers, competition from different sectors and continuously changing legislations have created serious pressure on banks. Today, customer satisfaction is more important key factor on being ahead of this highly competitive sector. To increase customer satisfaction, banks should improve with creative products and distribution channels to make differentiation in the powerful competition environment [1]. In order to solve the problems, banks try to find the customers who are interest of new products and services of the bank. Various machine learning techniques have been used to predict customer behaviour. Neural network is a computational learning system that uses a network of In order to solve the problems, banks try to find the customers who are interest of new products and services of the bank. Various machine learning techniques have been used to predict customer behaviour. Neural network is a computational learning system that uses a network of functions to understand and translate a data input of one form into a desired output. The aim of the study is to build a prediction model which is suitable to predict whether a customer is interest for long-term deposit or not. Neural network techniques such as Sequential Minimal Optimization, Support Vector Machine, Radial Basis Function Network, Multilayer Perceptron, Self Organizing Map, and Hierarchical Learning Vector Quantization have been used to build decision making model to predict customer response.

2. RELATED WORK

Moro et.al [2] in their proposed work implemented and compiled four different data mining algorithms namely, Decision Tree (DT), Logistic Regression (LR), Support Vector Machine (SVM), Neural Network (NN) to predict the success of telemarketing business strategies in banks. The results indicate that NN is the best algorithm with Area Under Curve (AUC) amount of 0.832. Elsalamony [3] has been comparing the performance of four different classification techniques such as multilayer perception neural network (MPLNN), tree augmented Naïve Bayes (TAN), logistic regression (LR) and Ross Quinlan new decision tree model (C5.0) on the bank direct marketing dataset in terms of accuracy, sensitivity and specificity. Experimental results have shown that C5.0 achieved better performance than other techniques. Apampa [4] in their proposed work compared the techniques Decision Tree, Nave Bayes and Logistic Regression and found that Decision Tree performed more better with AUC and CA value of 76.60 percent. Asare-Frempong et.al [5] in their proposed work apply four classification techniques namely, Multilayer Perceptron Neural Network (MLPNN), Decision Tree (C4.5), Logistic Regression and Random Forest (RF). Results showed that Random Forest Classifier with an accuracy of 87% is the better predictive ability among the four classifiers. Parlar et.al [6] focused on defining the relevant features for increasing

Bank Telemarketing effectiveness of customer subscription to term deposits. Chi-square and Information Gain were used in this regard. In what seemed like a mini evaluation, Precision and recall measures were used. They concluded that a reduced set of attributes improved classification performance.

3. METHODOLOGY

3.1 Sequential Minimal Optimization (SMO)

The SMO algorithm [7] is a specialized optimization approach for the SVM quadratic program. It takes advantage of the sparse nature of the support vector problem and the simple nature of the constraints in the Support Vector Machine Quadratic Programming (SVMQP) to reduce each optimization step to its minimum form. SMO breaks the large QP problem into a series of smallest possible QP problems. These small QP problems are solved analytically

3.2. Support Vector Machine (SVM)

The SVM [8] is a new method for binary classification and also classifies both nonlinear and linear data. SVM uses a nonlinear mapping to transform the original training data into a high-dimensional feature space and use a linear discriminator to classify the data value. Within the new dimension, it searches the decision boundary and separate the tuples of one class from another. SVM minimizes the expected error and finds the hyperplane using training tuples and margins.

3.3. Multilayer Perceptron

A multilayer perceptron [9] is a feed forward artificial neural network that generates a set of outputs from a set of inputs. Multilayer perceptrons (MLPs) are layered feedforward networks typically trained with static backpropagation. This network consists of three layers namely, input layer, hidden layer and output layers For a given set of training input-output pair, the algorithm provides a procedure for changing the weights in a back-propagation network to classify the given input patterns correctly. In this algorithm the weights are calculated during the learning period of the network. The training of the algorithm requires three stages: the feed-forward of the input training pattern, the calculation and back-propagation of the error, and update of weights. Once the network is trained it can produce its outputs very rapidly.

3.4. Self Organizing Map (SOM)

The Self-Organizing Map (SOM) [10] is a competitive network where the goal is to transform an input data set of arbitrary dimension to a one or two dimensional topological map. The SOM aims to discover underlying structure, e.g. feature map of the input data set by building a topology preserving map which describes neighbourhood relations of the points in the data set . SOM transforms a high dimensional input data domain to elements of a low dimensional array of nodes. The SOM array is essentially a fixed size grid of nodes. The training utilizes competitive learning, meaning that neuron with weight vector that is most similar to the input vector is adjusted towards the input vector and the neuron is said to be the “winning neuron” or the Best Matching Unit (BMU). The weights of the neurons close to the winning neuron are also adjusted but the magnitude of the change depends on the physical distance from the winning neuron which is decreased with time.

3.5. Radial Basis Function Network (RBFN)

Radial Basis function (RBF) network [11] is a nonlinear hybrid network which contains an input layer, a single hidden layer with a non-linear RBF activation function and an output layer. The input layer reads n inputs, the hidden layer consists of m radial basis functions and the output layer consisting of a linear additive function which produces the response. The input neurons are linear i.e they pass the input to the hidden neurons without any processing. In a RBF network, different layers perform different tasks. Therefore it is useful to separate the optimization of the hidden unit and output layers of the network by using different techniques. In this network, there are feed-forward connections between input and hidden layers and between hidden and output layers.. The input layer feeds forward to each hidden neuron. Using radial basis function the hidden neurons compute the signal and pass on these signals through weighted pathways to the linear output neuron which generates the output signal by summing up these input signals.

3.6. Hierarchical Learning Vector Quantization (HLVQ)

HLVQ [12] divides a feature space hierarchically by a few reference vector. HLVQ divides the feature space into regions by the overlapping technique. The regions divided at a layer is further divided at the lower layer hierarchically. During learning process neurons are arranged hierarchically like tree structure. The learning process continued until all neurons satisfying the terminating condition

4. THE PROPOSED MODEL

The proposed model consists of two major layers as depicted in figure 1. In the first layer irrelevant attributes are removed using three attribute evaluator methods namely, filtered attribute evaluator, one-R attribute evaluator, and relief-F attribute evaluator. In the second layer the reduced dataset is classified using six artificial neural network based techniques viz., Sequential Minimal Optimization(SMO), Support Vector Machine (SVM), Multilayer Perceptron (MP), Self-Organizing Map (SOM), Radial Basis Function Network(RBFN), and Hierarchical Learning Vector Quantization (HLVQ). Further, 10-fold cross validation technique is used for training and testing of the model. The performance of the model evaluate using certain standard criteria

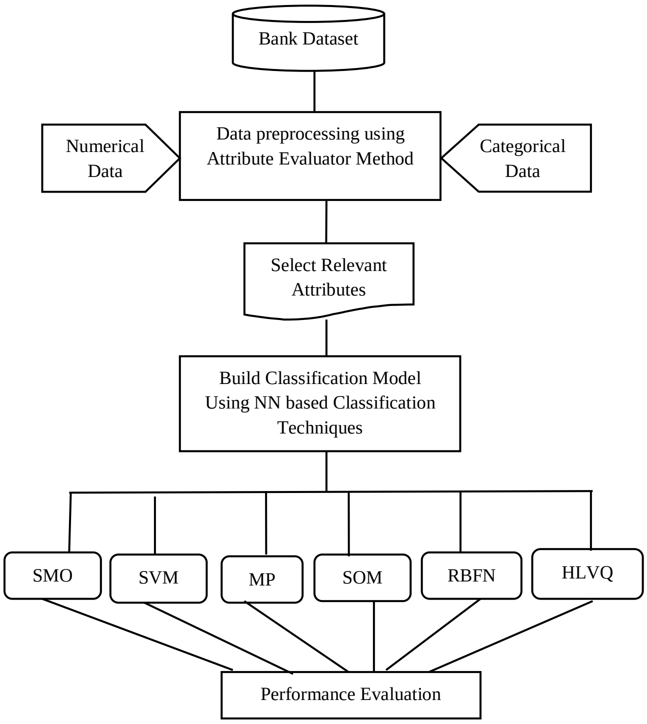

Fig.1 The proposed Model

5. EXPERIMENTAL SETUP

5.1. Bank Dataset

The dataset used for experiment is the bank marketing dataset from UCI Repository [13]. It includes 17 campaigns of a Portuguese bank conducted between May 2008 and November 2010. The dataset is related with direct marketing campaigns of a Portuguese banking institution. The dataset consists of three types of attributes: nominal, binary, and numerical features.

Table 1: Attribute description of Bank Dataset

#	Attributes	Category	Attribute Description
1	Age	Numeric	The age of the customer
2	Job	Categorical	Client's occupation
3	Martial Status	Categorical	Marital status
4	Education	Categorical	The education level
5.	Default	Binary	Has credit in default
6	Balance	Numeric	Average yearly balance, in euros
7	Housing	Binary	Has housing loan
8	Loan	Binary	Has personal loan
9	Contact	Categorical	Contact communication type
10	Day	Numeric	Last contact day of the month
11	Month	Numeric	Last contact month of year
12	Duration	Numeric	Last contact duration, in seconds
13	Campaign	Numeric	Last contact day of the month
14	Pdays	Numeric	Number of days that passed by after the client was last contacted from a previous campaign
15	Previous	Numeric	Number of contacts performed before this campaign and for this client
16	Poutcome	Categorical	Outcome of the previous marketing campaign
17	Output	Categorical	Has the client subscribed a term deposit

5.2. Attribute Selection

In problems dealing with high dimensional feature space, some of the attributes may be redundant or irrelevant. Removing these redundant or irrelevant attributes is very important for effectiveness. Attribute selection aims to identify small subset of the relevant attributes from the original features by removing redundant, irrelevant or noisy attributes of a dataset that could improve the performance of classification algorithms. Attribute selection can lead to better learning performance such as higher learning accuracy and lower computational cost. Here three attribute selection methods, viz., Filtered attribute evaluator, OneR attribute evaluator, and ReliefF attribute evaluator are applied for selection of important attributes.

Filters based feature selection [14] evaluate the usefulness of features in prediction independent of any learning algorithm. Filters are fast and are computationally more efficient but totally ignore the dependency of features' worthiness on learning algorithms.

One-R Attribute Evaluator [15] algorithm generates a one-level decision tree expressed in the form of a set of rules all of which test one particular attribute. It is capable of generating good rules for characterizing the structure in data. One-R can handle missing values and numeric attributes.

The One-R algorithm creates rules and tests a single attribute at a time and branch for every value of that attribute. For every branch, the class with the best classification is selected. To create a rule for an attribute, the most frequent class for each attribute value must be determined. Rule based algorithms follow three steps, viz., Generate rule R on training data S, remove the training data covered by rule, and repeat the process.

The three basic steps of Relief-F [16] attribute evaluator technique are:

1. Calculate the nearest miss and nearest hit.
2. Calculate the weight of a feature.
3. Return a ranked list of features or the top K features according to a given threshold

5.4. Performance Measurement

The performance of the classification model is evaluated using the statistical measures namely, accuracy and sensitivity. These measures are defined by a confusion matrix that contains information about actual class and predicted class. The confusion matrix is a table with two rows and two columns that reports the number of False Positive (FP), False Negative (FN), True Positive (TP), and True Negative (TN). The following table shows the confusion matrix for a two class classifier.

Table 2 : Confusion Matrix

		Predicted Class	
		Negative Class	Positive Class
Actual Class	Negative Class	True Negative (TN)	False Positive (FP)
	Positive Class	False Negative (FN)	True Positive (TP)

True Positive (TP) : Observation is positive and is predicted to be positive.

False Negative (FN) : Observation is positive and is predicted to be negative.

True Negative (TN): Observation is negative and is predicted to be negative.

False Positive (FP) : Observation is negative, but is predicted positive.

Accuracy measures the probability that an algorithm can correctly predict positive and negative examples. Accuracy is calculated as:

Accuracy = (TP + TN) / (TN + TP + FN + FP)

Specificity measures the probability that an algorithm can correctly predict negative examples. It is also called true negative rate (TNR). Specificity is calculated as

Specificity = (TN) / (TN + FP)

6. RESULT ANALYSIS

Six different neural network based classifiers, namely, Sequential Minimal Optimization, Support Vector Machine, Radial Basis Function Network, Multilayer Perceptron, Self Organizing Map, and Hierarchical Learning Vector Quantization with three attribute evaluator methods were applied on the bank dataset. The performance of different classifiers are evaluated on the basis of accuracy and specificity. The result shows that Multilayer Perceptron technique with filtered attribute evaluator gives highest accuracy of **90.0179%** and Support Vector Machine technique with Relief-F attribute evaluator gives highest specificity of **99.9374 %.** Accuracy and specificity of different combinations of classifiers with attribute evaluators are presented in figures 2, 3, 4, 5, 6, and 7 respectively.

Table 3 : Comparision of Classifiers with Feature Reduction

Feature Selection Techniques	Classifier Techniques	Evaluation Criteria	
		Accuracy in %	Specificity in %
Filtered Attribute Evaluator	SMO	89.2858	88.3015
	SVM	78.273	87.7636
	Multilayer Perceptron	**90.0179**	95.8168
	SOM	88.286	**98.8903**
	RBFN	89.0093	96.661
	Hierarchical LVQ	88.2307	86.041
OneR Attribute Evaluator	SMO	89.2858	98.6649
	SVM	88.14	**99.5792**
	Multilayer Perceptron	89.3676	95.1405
	SOM	88.9164	97.7005
	RBFN	**89.5645**	96.691
	Hierarchical LVQ	88.7262	96.9415
RliefF Attribute Evaluator	SMO	89.2858	98.6649
	SVM	88.2617	**99.9374**
	Multilayer Perceptron	**89.7348**	95.2332
	SOM	88.9385	98.2816
	RBFN	89.4893	96.8914
	Hierarchical LVQ	88.3856	97.2997

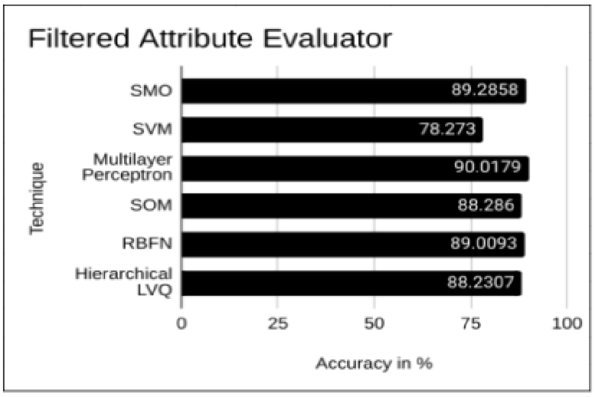

Fig. 2 Comparison of Accuracy

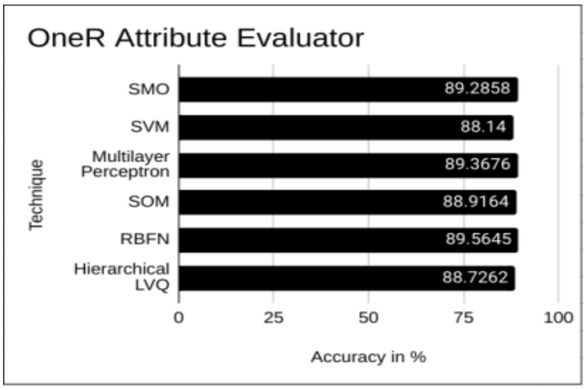

Fig. 3 Comparison of Accuracy

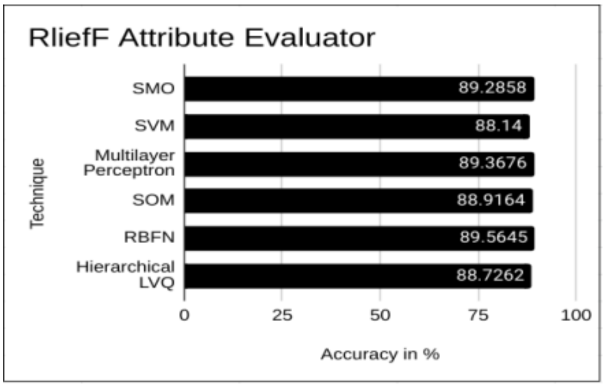

Fig. 4 Comparison of Accuracy

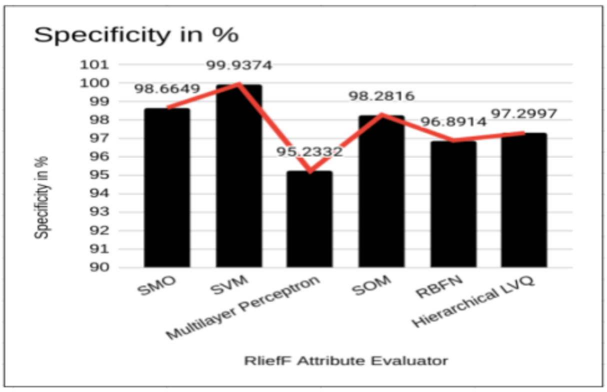

Fig. 5 Comparison of Specificity

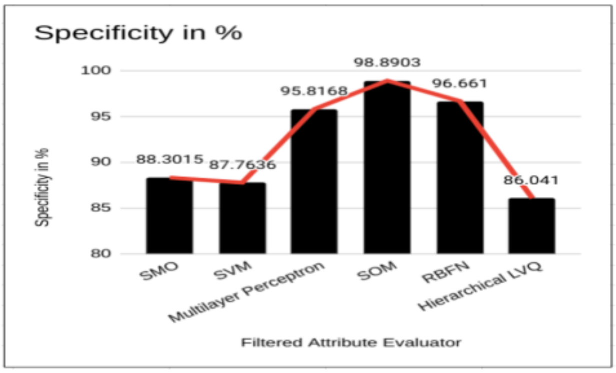

Fig. 6 Comparison of Specificity

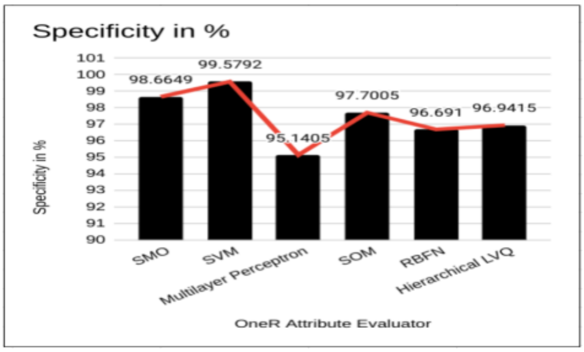

Fig. 7 Comparison of Specificity

7. CONCLUSION

This paper have been evaluated and compared the performance of six different classification neural network techniques namely, SMO, SVM, RBFN, MP, SOM, and HLVQ on the bank direct marketing dataset to predict long-term deposit interest of customers in the bank. The classification performances of the models are analyzed using the evaluation criteria such as accuracy and specificity. It was observed that Support Vector Machine technique with Relief-F attribute evaluator gives highest specificity.

REFERENCES

[1] S. Karaslar, "Bankacılık Sektöründe Pazarlama Uygulamaları: Garanti Bankası Örneği ". Trakya Üniversities Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı Yüksek Lisans Dönem Projesi , 2014.

[2] S. Moro, P. Cortez, and R. Laureano . A data mining approach for bank telemarketing using the rminer package and r tool , 2013.

[3] Hany A. Elsalamony , " Bank Direct Marketing Analysis of Data Mining Techniques " , International Journal of Computer Applications , vol. 85, no. 7, pp. 12-22, 2014.

[4] O.Apampa, "Evaluation of classification and ensemble algorithms for bank customer marketing response prediction", Journal of International Technology and

Information Management , vol 25, no.4, pp. 85-100, 2016.

[5] Justice Asare-Frempong and Manoj Jayabalan (2017). “Predicting Customer Response to Bank Direct Telemarketing Campaign”, The International Conference on Engineering Technologies and Technopreneurship (ICE2T 2017, IEEE) , 2017.

[6] T.Parlar and S. K. Acaravci, “Using Data Mining Techniques for Detecting the Important Features of the Bank Direct Marketing Data”, International Journal of Economics and Financial Issues, vol. 7, no. 2, pp.692–696, 2017.

[7] J. Platt, Fast Training of Support Vector Machines using Sequential Minimal Optimization, in Advances in Kernel Methods – Support Vector Learning, MIT Press , 1998.

[8] J. Han and M. Kamber, “Data Mining Concepts and Techniques” , Morgan Kaugmann Publishers, 2008.

[9] S.N. Sivanandam, S.N. Deepa, Principles of Soft Computing, Wiley-India, 2011.

[10] T.Kohonen, “ The Self-Organizing Map” , In the Proceedings of the IEEE, Vol.78, Issue : 9, pp. 1464-1480, 1990.

[11] S.V. Chakravarthy and J. Ghosh, “Scale Based Clustering using Radial Basis Function networks”, Proceeding of IEEE International Conference on Neural Networks, pp.897-902, 1994.

[12] T. Kohonen, “’Improved version of learning vector quantization” Proc. International Joint Conference on Neural Networks, pp.I 545–550, San Diego, June 1990.

[13] http://archive.ics.uci.edu/ml (The UCI Machine Learning Repository is a collection of databases)

[14] N. Abe and M. Kudo, “Entropy criterion for classifier-independent feature selection”, *Lecture Notes in Computer Science*, Vol. 3684, pp. 689-695, 2005

[15] R. C. Holte, “Very simple classification rules perform well on most commonly used datasets”, Machine Learning, p. 63-91,1993.

[16] R. S. Marko and K. Igor, “Theoretical and empirical analysis of relief and reliefF”, Machine Learning Journal, pp. 23-69, 2003.