Proceedings of the World Congress on Engineering 2014 Vol I,
WCE 2014, July 2 - 4, 2014, London, U.K.

Customer Behavior on RFMT Model Using Neural Networks

Qeethara Kadhim Al-Shayea, *Member, IAENG*, and Tamara Kadhim Al-Shayea

***Abstract*—Customer data is critical to marketing success. The goal of this study is to predict customer behavior using a supervised learning neural network. Feed-forward back propagation network with tan-sigmoid transfer functions is used as a classifier to predict whether a customer will buy in this month or not. Scaled conjugate gradient (SCG) algorithm is used with proposed neural network. This algorithm combines the model-trust region approach with the conjugate gradient approach. The results of applying the proposed artificial neural networks methodology to predict based upon recency, frequency, monetary, and time (RFMT) model show abilities of the network to learn the patterns corresponding to RFMT of the customer. The data set is obtained from UCI machine learning repository. The percent correctly classified in the simulation sample by the proposed neural network is 89 percent.**

***Index Terms*—Artificial Neural Networks, Feed-forward back propagation network, REMT parameters, Customer behavior prediction and Artificial Intelligence.**

I. Introduction

THE prediction problem is vital topics in many areas. Many problems in business, science, industry, and medicine can be treated as prediction problems. Examples include bankruptcy prediction, credit scoring, medical diagnosis, and customer behavioral. RFMT analysis using artificial neural networks (ANNs) is a marketing technique used for analyzing customer behavior. ANN have been applied to a variety of business areas such as accounting, auditing, finance, management, decision making, marketing and production.

Cho et al. [1] proposed a new incremental weighted mining based on RFM analysis for recommending prediction in u-commerce. The proposing method can extract frequent items and create weighted association rules using incremental weighted mining based on RFM analysis rapidly when new data are added persistently in order to predict frequently changing trends by emphasizing the important items with high purchasability according to the threshold for creative weighted association rules in u-commerce.

Al-Shayea and El-Refae [2] presented a network to predict customer behavior using neural network. The results of applying the artificial neural networks methodology to predict based upon recency, frequency, monetary, and time (RFMT) model show abilities of the network to learn the patterns corresponding to RFMT of the customer.

Manuscript received 8 April, 2014.

Q. K. Al-Shayea is Associate Professor with the Department of Management Information Systems, Al-Zaytoonah University of Jordan, Amman, Jordan (e-mail: kit_alshayeh@yahoo.com).

T. K. Al-Shayea is a programmer in Warka Company for Telecommunications. (e-mail: tamara.kazim@yahoo.com).

Liou and Tzeng [3] used the Dominance-based Rough Set Approach (DRSA) to provide a set of rules for determining customer attitudes and loyalties, which can help managers, develop strategies to acquire new customers and retain highly valued ones. A set of rules is derived from a large sample of international airline customers, and its predictive ability is evaluated. The results, as compared with those of multiple discriminate analyses, are very encouraging. They prove the usefulness of the proposed method in predicting the behavior of airline customers.

Yeh, Yang and Ting [4] introduced a comprehensive methodology to discover the knowledge for selecting targets for direct marketing from a database. This study expanded RFM model by including two parameters, time since first purchase and churn probability. Using Bernoulli sequence in probability theory, they derive out the formula that can estimate the probability that one customer will buy at the next time, and the expected value of the total number of times that the customer will buy in the future.

Cheng and Chen [5] proposed a new procedure, joining quantitative value of RFM attributes and K-means algorithm into rough set theory (RS theory), to extract meaning rules, and it can effectively improve these drawbacks.

Sant'Anna and Ribeiro [6] presented a new form of directly combining the values of the RFM variables is compared to an approach based on fitting a stochastic model. This last model is a mixture of a model for the number of transactions and another for the value spent.

Chiu et al. [7] proposed a market segmentation system based on the structure of decision support system which integrates conventional statistic analysis method and intelligent clustering methods such as artificial neural network, and particle swarm optimization methods. The proposed system is expected to provide precise market segmentation for marketing strategy decision making and extended application.

Chen et al. [8] incorporated the recency, frequency, and monetary (RFM) concept presented in the marketing literature to define the RFM sequential pattern and develop a novel algorithm for generating all RFM sequential patterns from customers' purchasing data. Using the algorithm, they proposed a pattern segmentation framework to generate valuable information on customer purchasing behavior for managerial decision-making. Extensive experiments are carried out, using synthetic datasets and a transactional dataset collected by a retail chain in Taiwan, to evaluate the proposed algorithm and empirically demonstrate the benefits of using RFM sequential patterns in analyzing customers' purchasing data.

Chan [9] presented a novel approach that combines customer targeting and customer segmentation for

campaign strategies. Additionally, he proposed using generic algorithm (GA) to select more appropriate customers for each campaign strategy.
Li, Shue and Lee [10] proposed a Business Intelligence process for Internet Service Providers (ISP) dealers in Taiwan to assist management in developing effective service management strategies.

II. Artificial Neural Networks

An artificial neural networks ANN model emulates a biological neural network. Neural computing actually uses a very limited set of concepts from biological neural systems. It is more of an analogy to the human brain than an accurate model of it. Neural concepts are usually implemented as software simulations of the massively parallel processes that involve processing elements (also called artificial neurons) interconnected in network architecture. The artificial neuron receives inputs analogous to the electrochemical impulses the dendrites of biological neurons receive from other neurons. The output of the artificial neuron corresponds to signals sent out from a biological neuron over its axon. These artificial signals can be changed by weights in a manner similar to the physical changes that occur in the synapses as shown in Fig. 1.

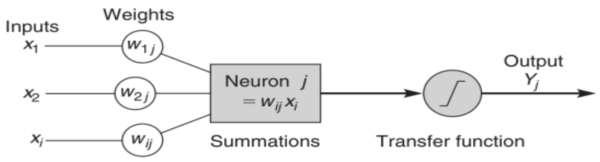

Fig. 1 Processing information in an artificial neuron

An artificial neural network (ANN) is a computational model that attempts to account for the parallel nature of the human brain. An (ANN) is a network of highly interconnecting processing elements (neurons) operating in parallel. These elements are inspired by biological nervous systems. As in nature, the connections between elements largely determine the network function. A subgroup of processing element is called a layer in the network. The first layer is the input layer and the last layer is the output layer. Between the input and output layer, there may be additional layer(s) of units, called hidden layer(s). Neural network can be train to perform a particular function by adjusting the values of the connections (weights) between elements. Fig. 2 represents a typical neural network with one hidden layer.

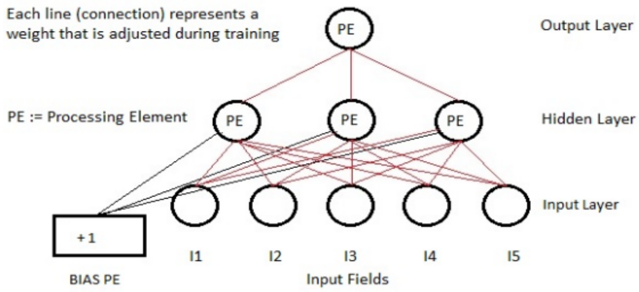

Fig. 2 A typical neural networks with one hidden layer

III. The Proposed Prediction Model

The recency, frequency, monetary and time (RFMT) method is an approach used to measure costumer's loyalty and segment customers into various group for future personalization services. This study identifies customer behavior using (RFMT) model**.**

An ANN have been shown to be very promising systems in many forecasting applications and business classification applications due to their ability to “learn” from the data, their non parametric nature and their ability to generalize. A feed-forward back propagation neural network with tan-sigmoid transfer functions in both the hidden layer and the output layer is proposed to predict the customer behavior. It consists of three layers: the input layer, a hidden layer, and the output layer. A one hidden with 20 hidden layer neurons is created and trained. The proposed network uses the scaled conjugate gradient algorithm for training. The input and target vectors are automatically divided into three sets: 60% are used for training, 20% are used to validate that the network is generalizing and to stop training before over fitting, and the last 20% are used as a completely independent test of network generalization. The training set is used to teach the network. Training continues as long as the network continues improving on the validation set. The test set provides a completely independent measure of network accuracy. The information moves in only one direction, forward, from the input nodes, through the hidden nodes and to the output nodes. There are no cycles or loops in the network. The proposed neural networks are shown in Fig.3.

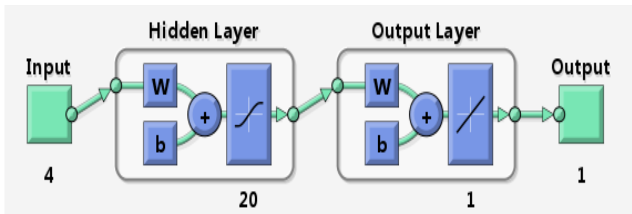

Fig. 3 The proposed network

The hidden neurons in feed-forward neural network are able to learn the pattern in data during the training phase and mapping the relationship between input and output pairs. Each neuron in the hidden layer uses a transfer function to process data it receives from input layer and then transfers the processed information to the output neurons for further processing using a transfer function in each neuron. Tan-sigmoid transfer functions in both the hidden layer and the output layer is used.
The output of the hidden layer can be represented by

$$Y_{Nx1} = f(W_{NxM} X_{M,1} + b_{N,1}) \quad (1)$$

where Y is a vector containing the output from each of the N neurons in a given layer, W is a matrix containing the weights for each of the M inputs for all N neurons, X is a vector containing the inputs, b is a vector containing the biases and $f(\cdot)$ is the activation function [11].

ISBN: 978-988-19252-7-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

IV. EXPERMENTAL RESULTS

A. *Data Analysis*

The data is obtained from UCI Machine Learning Repository. The data was adopted from Blood Transfusion Service Center in Hsin-Chu City in Taiwan. The center passes their blood transfusion service bus to one university in Hsin-Chu City to gather blood donated about every three months. The main idea of this data set is to construct the neural network model, which will perform the presumptive prediction of customer behavior. These donor data, each one included R (Recency - months since last donation), F (Frequency - total number of donation), M (Monetary - total blood donated in c.c.), T (Time - months since first donation), and a binary variable representing whether he/she donated blood in March 2007 (1 stand for donating blood; 0 stands for not donating blood).

The data set contains 748 samples. 500 sample used in training the network while 248 samples used in testing the network. Table-I presents the parameters which are considered as predictor variables. The data set is processed using Microsoft Office Excel to be appropriate to the neural network.

TABLE I
PREDICTOR VARIABLE OF DATASETS USED IN THE STUDY

Blood Transfusion Service Center data	
S. No.	Predictor Variable Name
1	R (Recency - months since last donation)
2	F (Frequency - total number of donation)
3	M (Monetary - total blood donated in c.c.)
4	T (Time - months since first donation)

B. *Performance Evaluation*

A two-layer feed-forward network with 4 inputs and 20 sigmoid hidden neurons and linear output neurons was created as shown in Fig. 3. The dataset contains 748 samples. 500 sample used in training the network while 248 samples used in testing the network. Training is done using scaled conjugate gradient back propagation network. The scaled conjugate gradient algorithm (SCG) developed by Moller [12] was designed to avoid the time-consuming line search. This algorithm combines the model-trust region approach with the conjugate gradient approach.

The results of applying the proposed neural network to distinguish between buyer and non buyer customer based upon (RFMT) model showed very good abilities of the network to learn the patterns corresponding to (RFMT) parameters. The results were good; the network was able to classify approximately 88% of the cases in the training set as shown in Fig. 4. Fig. 4 shows the confusion matrices for training, testing, and validation, and the three kinds of data combined. The diagonal cells in each table show the number of cases that were correctly classified, and the off-diagonal cells show the misclassified cases. The blue cell in the bottom right shows the total percent of correctly classified cases (in green) and the total percent of misclassified cases (in red). Fig.5 shows the training state values. Best validation performance is 0.10389 at epoch 10 as shown in Fig.6. The mean squared error (MSE) is the average squared difference between outputs and targets.

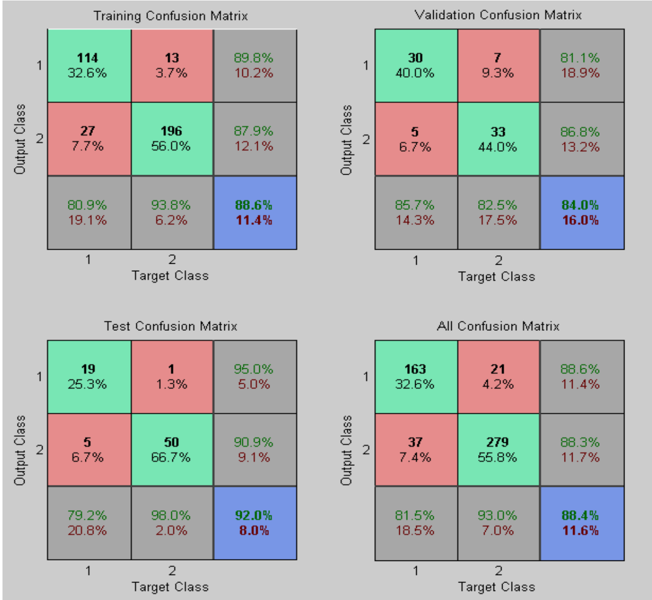

Fig. 4. The confusion matrices

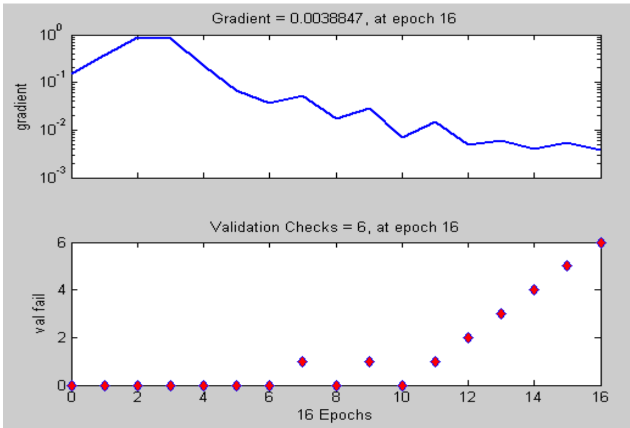

Fig. 5. The training state values

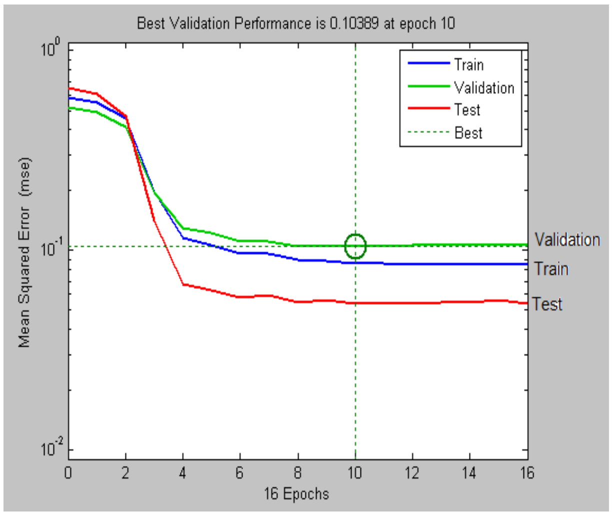

Fig. 6. The proposed network performance

ISBN: 978-988-19252-7-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

The percent correctly classified in the simulation sample by the feed-forward back propagation network is approximately 89 percent as shown in Fig.7.

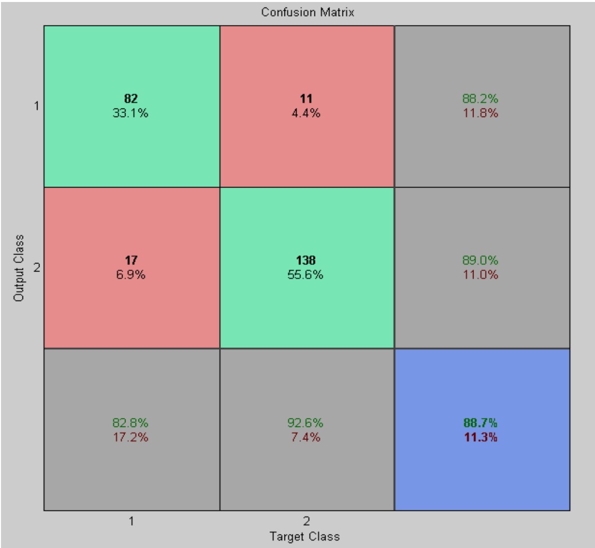

Fig. 7. The confusion matrix

V. CONCLUSION

Artificial neural networks have powerful pattern classification and prediction capabilities. In this study feed-forward back propagation neural network with tan-sigmoid transfer functions in both the hidden and the output layer is applied for predict the customer behavior based upon the RFMT model. Experiments show that SCG give significant results in prediction of the customer behavior.

REFERENCES

[1] Y. S. Cho, S. C. Moon, I. Oh, J. Shin and K. H. Ryu, Incremental Weighted Mining based on RFM Analysis for Recommending Prediction in u-Commerce, International Journal of Smart Home Vol.7, No.6, 2013, pp.133-144.

[2] Q. K. Al-Shayea and G. A. El-Refae, Artificial Neural Networks for Customer Behavior on RFMT Model, Business Research Yearbook Balancing Profitability and Sustainability: Shaping the Future of Business Vol. XVIII, No. 1, 2011, pp. 103-108.

[3] J. H. Liou and G. Tzeng, A Dominance-based Rough Set Approach to customer behavior in the airline market, Information Sciences 180, 2010, pp. 2230–2238.

[4] I. Yeh, K. Yang and T. Ting, Knowledge discovery on RFM model using Bernoulli sequence, Expert Systems with Applications, Vol. 36, No. 3, 2009, pp. 5866-5871.

[5] C. Cheng, and Y. Chen, Classifying the segmentation of customer value via RFM model and RS theory, Expert Systems with Applications, Vol. 36, No. 3, 2009, pp. 4176-4184.

[6] A. P. Sant'Anna and R. O. Ribeiro, Statistical modeling and probabilistic composition in the prediction of the customer lifetime value, Benchmarking: An International Journal, Vol. 16, No. 3, 2009, pp. 335–350.

[7] C. Chiu, Y. Chen, I. Kuo and H. Ku, An intelligent market segmentation system using k-means and particle swarm optimization, Expert Systems with Applications, Vol. 36, No. 3, 2009, pp.4558-4565.

[8] Y. Chen, M. Kuo, S. Wu, and K. Tang, Discovering recency, frequency, and monetary (RFM) sequential patterns from customers' purchasing data, Electronic Commerce Research and Applications, Vol.8, Issue 5, 2009, pp. 241-251.

[9] C. C. H. Chan, Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer, Expert Systems with Applications, Vol. 34, No. 4, 2008, pp. 2754-2762.

[10] S. Li, L. Shue, and S. Lee, Business intelligence approach to supporting strategy-making of ISP service management, Expert Systems with Applications, Vol. 34, No. 3, 2008, pp. 739-754.

[11] J. A. Freeman, and D. M. Skapura, Neural networks: algorithms, applications and programming techniques, Addison Wesley Longman, 1991.

[12] M. F. Moller, A scaled conjugate gradient algorithm for fast supervised learning, Neural Networks, Vol. 6, 1993, pp. 525-533.