Industrial Engineering
& Management Systems
Vol 15, No 4, December 2016, pp.324-334
ISSN 1598-7248 │ EISSN 2234-6473 │
https://doi.org/10.7232/iems.2016.15.4.324

Application of a Hybrid System of Probabilistic Neural Networks and Artificial Bee Colony Algorithm for Prediction of Brand Share in the Market

Jamal Shahrabi*, Sara Mottaghi Khameneh
Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran

(Received: November 28, 2016 / Revised: December 6, 2016 / Accepted: December 12, 2016)

ABSTRACT

Manufacturers and retailers are interested in how prices, promotions, discounts and other marketing variables can influence the sales and shares of the products that they produce or sell. Therefore, many models have been developed to predict the brand share. Since the customer choice models are usually used to predict the market share, here we use hybrid model of Probabilistic Neural Network and Artificial Bee colony Algorithm (PNN-ABC) that we have introduced to model consumer choice to predict brand share. The evaluation process is carried out using the same data set that we have used for modeling individual consumer choices in a retail coffee market. Then, to show good performance of this model we compare it with Artificial Neural Network with one hidden layer, Artificial Neural Network with two hidden layer, Artificial Neural Network trained with genetic algorithms (ANN-GA), and Probabilistic Neural Network. The evaluated results show that the offered model is outperforms better than other previous models, so it can be use as an effective tool for modeling consumer choice and predicting market share.

Keywords: Consumer Choice Model, Brand Share, Artificial Neural Network, Modeling, Predicting, Probabilistic Neural Network, Artificial Bee Colony Algorithm

* Corresponding Author, E-mail: jamalshahrabi@aut.ac.ir

1. INTRODUCTION

Market share, which is usually defined as the ratio of selling business to sales ratio in the industry, is one of the variables of market structure, and it has been a subject of important studies in the field of economy in recent decades (Barthwal, 2007).

Nowadays, in all around the world companies are trying to gain more market share among their competitors to increase their ratio of sales. In order to achieve competitive advantage, companies use their own resources to improve their relative position toward their competitors. Improving company's position in market can lead to obtain better opportunities for companies in future. In studies of industrial organization, market share and growth have been firmed as key performance indicators and positive relationship between market share and economic profit has been proved (Hansen and Wernerfelt, 1989).

The level of market share indicates the carrier's degree of monopoly power. High market share provides higher monopoly power, while low market share leads to little or none (Borenstein, 1990). If market share increases, a company is likely to have more profit margins, a decrease in marketing costs higher quality (Buzzell *et al*., 1975).

Empirical evidences reveal that consumer choice models can be use as a market response simulator, so they can be use for market share predicting (Fish *et al*., 2004). Identification and modeling of consumer choice has always been one of the most important branches of Marketing Research (van Wezel and Potharst, 2007).

Consumer choices models are model the choice of different brands by customers in the market. These models can be applied as simulators of market response to the marketing strategies of firms in order to suggest appropriate alternative decisions on price promoting, policies of companies, and so on. Thus, development of a proper model with high accuracy on the prediction of customer behavior in the market can be of help to companies as a strategic tool to improve the adoption of decision-making policies in the scope of marketing (Banerjee *et al*., 2005).

Therefore a reliable and convenient brand choice model acts as a market response simulator, and Managers who use them are able to simulate marketing decisions and improve the efficiency of their decisions. A market response simulator based on the customer's choice models, enables marketing managers to deliberate hundreds of "what-if" Pricing and strategies scenarios. For example, one can answer "what-if" questions like: what happens to my brand's market share if the price increases, if I add a new product line in my business, or if I increase promotional activities, or if a competitor decreases its price? For an in-depth discussion of benefits of marketing response models, the reader is respectively referred to Wiernga *et al*. (2008).

Studies show that the superiority of Artificial Neural Network in predicting brand share in comparison with Multi Nominal Logit (MNL) first shown by Agrawal and Schorling (1996). Result of This study led to be ANNs become accepted tools for market researchers (Fish *et al*., 2004; Agrawal and Schorling, 1996). Nowadays because of the access to market response data such as individual customer purchases, price, discounts, promotion, distribution, marketing, commercial strategies, and advertising there is an increasing interest among companies for using market response models (Fish *et al*., 2004).

ANN has been widely used in modeling brand choice and customer behavior. Kumar *et al*. (1995) compared ANN and regression models in predicting consumer choice. West *et al*. done a comparative analysis of ANNs and statistical methods for predicting consumer choice and they explored advantages and disadvantages of ANN compared to statiscal methods (West *et al*., 1997). Bentz and Merunka developed a hybrid model in which ANN and MNL combine together in the brand choice modeling (Bentz and Merunka, 2000). Hruschka *et al*. (2002), used feed-forward ANN for discovering nonlinear effects on brands' utilities and compared the performance of this model with different MNL models. Hu and Tsoukalas (2003) used ANN models and the ensemble technique of stacsked generalization to investigate the relative importance of situational and demographic factors on consumer choice. Vroomen *et al*. (2004) proposed a two step ANN choice modeling framework in the first step of which they took consideration sets of the households into account where the consideration set corresponds with the hidden layer of the network. They have illustrated their model for the choice between six detergent brands and showed that the model improves upon one-step models. Hruschka (2007) used a Heterogeneous Multinomial Probit (MNP) Model with ANN extension to model brand choice. Hu *et al*. (2008) have shown how ANN can be used to model consumer choice. Their study focuses on two key issues in ANN modeling, model building and feature selection. Lee *et al*. (2008) applied neural network to classify consumers' behavior in choosing hospitals. Their results showed that neural network model is useful in identifying existing patterns of hospitals' consumers. Kaya *et al*. (2010) compared the performances of ANN and MNP approaches in modeling the choice decision, they results showed that ANN's predictions are better while MNP is useful in providing marketing insight.

The accuracy of predictions is the most important factor in the selection of forecasting models. Using hybrid models or combination of different models is a common way to increasing accuracy of predictions. Literature related to hybrid models is widespread and several studies have been done since the first researches done by Reid (1968), Bates and Granger (1969). Clemen (1989) has done an overview in this regard. Experimental and theoretical results have shown that the combination of different models is an effective and efficient way to improve the accuracy of predictions (Khashei *et al*., 2008; Armano *et al*., 2005; Chen *et al*., 2007).

In predicting studies, many researchers have proposed hybrid models by using ANNs. Fish *et al*. (2004) introduced a new architectural approach to ANN choice modeling and used a feed-forward ANN trained with a Genetic Algorithm to model individual consumer choices and brand share in a retail coffee market. Recently, Probabilistic Neural Network (PNN) has been used in many studies successfully for solving classification problems to avoid deficiency of feed-forward ANNs trained by Back Propagation (BP) (Cheng *et al*., 2010). PNN is one of the ANNs which use a kernel-based approximation to form an estimate of the probability density function of categories in a classification problem and an algorithm for approximating the Bayesian decision rule (Specht, 1990). Wasserman suggests a few advantages for the PNN over the conventional BP neural networks as follows:

- The PNN does not require a separate training phase and is therefore computationally on average more than five times faster than the conventional BP networks.
- Additional training instances can be incorporated easily into a PNN as they become available in the future.
- The PNN provides classification robustness in domains with noisy data (Wasserman, 1993).

In the literature of brand choice problem, there are a few studies have used PNN in customer choice problem. Gan *et al*. (2005) had done a comparison between Probabilistic Neural Network, feed-forward ANN, and logistic model. They have shown that PNN outperforms better, in higher accuracy and speed, than both ANN and Logistic model in customer choice model. Kazemi *et al*. (2013) used a PNN trained by Dynamic Decay Adjustment Algorithm (DDA) to model customer choice. The evaluation process is carried out for modeling individual consumer choices in a retail coffee market. The evaluation results have been shown that the offered approach outperforms ANN and ANN trained with GA so it can be considered as an effective tool for consumer behavior modeling and simulation.

Artificial Bee Colony (ABC) algorithm is the one of the swarm intelligence which has been most widely studied and applied to solve the real-world problems. Karaboga and Basturk (2007) have shown better performance of ABC compared to GA, Particle Swarm Optimization (PSO), and particle swarm inspired evolutionary algorithm. Indeed, bee colony algorithm was found to have better performance than other algorithms. The advantages of using ABC in comparison to other approaches include the enjoyment of memory, local searches, and solution mechanism improvement and, thereby, this algorithm can be effective in finding the optimal solution with high quality. Bee colony algorithm can be optimized to provide a trade-off between performance and complexity of optimization (Karaboga and Basturk, 2007; Karaboga and Akay, 2009). ABC has been used in many applications in several different fields. One of the most interesting application areas is training neural networks. First of all Karaboga and Akay (2007) employed ABC for training feed-forward neural networks, i.e. searching optimal weight set. For reading more applications of ABC in training ANNs readers are referred to (Karaboga *et al*., 2014; Gorunescu, 2006).

In this study a new hybrid approach, that we were proposed to predict customer choice in the market, which uses ABC to find the optimal standard deviation in PNN applied to predict brand share of the companies. In comparison with hybrid neural network models with genetic algorithms, the proposed model has shown such advantages such as higher accuracy, insensitivity to outliers, a smaller number of control parameters. Since customer choice models can use as a market simulator, a case study is undertaken in order to test the proposed model and measure its modeling and predictive power in brand choice problem.

In this study, the training dataset will be held constant at the first stage of the “choice model”, but the network design changes from the standard model of neural network with back propagation training into a hybrid of probabilistic neural network and bee colony algorithm. All the networks are optimized to predict brand share.

In fact, the training datasets and abandoned data were held constant. All the models were optimized using mean absolute error of the predicted market share. Then, all the networks were evaluated on the abandoned sample to determine the predictive power of market share in terms of different coffee brands and sizes.

In the second stage, price sensitivity analysis was performed to demonstrate the usability and applicability of the PNN-ABC model.

2. METHODOLOGY

In this paper we use the model that we have proposed for customer choice modeling, PNN-ABC, to determine company’s brand share. Therefore, first we introduce PNN, ABC, and PNN-ABC.

2.1 The Probabilistic Neural Network

The Probabilistic Neural network is a special type of Neural Networks which using a kernel-based approximation to form an estimate of the probability density function of categories in a classification problem and an algorithm for approximating the Bayesian decision rule (Baghchesaraei *et al*., 2014; Specht, 1990).

PNNs consist of three layers of nodes. Since their training occurs in a path instead of a few numerical values per training vector, they are quickly learned. Probabilistic neural networks estimate the probability density functions of each category based on training samples. Figure 1 shows the structure of PNNs, which identify K classes. Probabilistic neural network can be extended to any number of categories.

The input layer contains N nodes: there is one node for each input feature of the input vector. Each feature input node is connected to all nodes in the latent layer; therefore, all the latent nodes fully receive the eigenvector.

Hidden nodes are divided into groups, that is, one group for each of the K categories. These groups are associated with eigenvectors in k-th class (here, there is a Gaussian function for each eigenvector). All Gaussian functions in the same category send the category of their values to one node in the output layer. Thus, there will be the number of K output nodes.

In the output node for class K, all Gaussian values of category K gather together and the total is scaled. Since the magnitude of probability is unique under the total function, the total appears as the probability density function. Here special symbols are used for clarity.

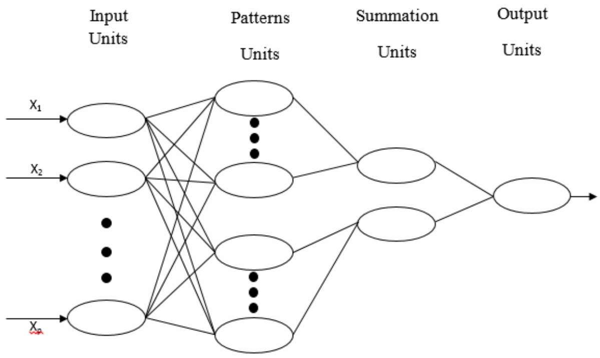

Figure 1. Probabilistic neural network structure.

P represents the eigenvectors of category 1, i.e. $\{x(p): p=1, \cdots, P\}$ and Q represents the eigenvectors of category 2, i.e. $\{x(q): q=1, \cdots, Q\}$. In the hidden layer, node P is placed in category 1 and Q lies in category 2.

Gaussian equations focused on the points of both categories 1 and 2, x (p) and y (q) are shown in vectors 2-1 and 2-2 for each input vector where N denotes the number of vectors:

$$g_1(x)=\left[1/\sqrt{(2\pi\sigma^2)^N}\right]\exp\left\{-\left\|x-x^{(p)}\right\|^2/(2\sigma^2)\right\} \quad (1)$$

$$g_2(y)=\left[1/\sqrt{(2\pi\sigma^2)^N}\right]\exp\left\{-\left\|y-y^{(q)}\right\|^2/(2\sigma^2)\right\} \quad (2)$$

σ values can be considered as half the size of the mean distances among the eigenvectors in the same group or they can be considered as half the distance of the sample with the nearest sample vector in each sample. The sum of the K-th node, which is obtained from the hidden nodes in the K-th group, is referred to as Parzen and/or Gaussian windows. The sums are defined through Eq. (3) and Eq. (4) as follows:

$$f_1(x)=\left[1/\sqrt{(2\pi\sigma^2)^N}\right]1/P\sum\nolimits_{p=1,p}\exp\left\{-\left\|x-x^{(p)}\right\|^2/(2\sigma^2)\right\} \quad (3)$$

$$f_2(x)=\left[1/\sqrt{(2\pi\sigma^2)^N}\right]1/Q\sum\nolimits_{q=1,Q}\exp\left\{-\left\|y-y^{(p)}\right\|^2/(2\sigma^2)\right\} \quad (4)$$

Here, x is the input eigenvector; σ represents extension parameters (standard deviation) for Gaussian function in classes 1 and 2. Subsequently, N represents the number of input vectors; P denotes the number of central vectors in class 1, Q shows the number of centers in class 2, and x (p) and y (q) represent centers of classes 1 and 2, respectively. $\left(\left\|x-x^{(p)}\right\|^2\right)$ is the Euclidean distance between x and x (p). Each input vector x is placed in both functions of $J_1(x)$, $f_2(x)$, and the maximum values of the two functions determine the output class. For problems with more than two classes, the same process is at play. Here there are no repetition and weight calculations (Baghchesaraei *et al*., 2015). For more information about probabilistic neural networks, readers are referring to SPECT 1990 (Baghchesaraei *et al*., 2015).

2.2 Artificial Bee Colony Algorithm

The ABC algorithm was introduced in 2005 by Karaboga. This algorithm is based on the collective behavior of bees in nature. In some studies have shown that bee colony performance is better than the other algorithms such as Genetic Algorithm, Particle swarm optimization, and other evolutionary algorithms. Because of simplicity in writing codes and reducing the need of parameter adjustment, ABC has been used in most research issues such as optimization functions, clustering analysis, and image processing, routing, and other issues (Kıran and Fındık, 2015).

This algorithm, with an initial population of random solutions, begins the search. Then by using the repeated process, tries to improve the random answers. According to the Eq. (5) initial population is created.

$$X_{i,j}=X_j^{min}+\Phi_{ij}\left(X_j^{max}-X_j^{min}\right) \quad (5)$$

In Eq. (5) Φ factor represents a random value between [-1, 1], $j\in\{1, 2, \cdots, D\}$, $i\in\{1, 2, \cdots, SN\}$, D is dimension or initial population size, and SN is the number of problem's parameters. X^{min}, X^{max} are the lowest and the highest limit of parameters values. Artificial bee colony consists of three groups of bees; employee bees, Scout bees, and onlooker bees. The number of employee bees is equal to the number of onlooker bees and this number is equal to the number of food sources or solutions; it means any employee bees are responsible for one solution in memory.

In the phase of employee bees, each employee seeks for a new solution in the neighborhood of solutions which are existence in the memory. It means that for each existence solution (x_i) a new neighborhood (v_i) according to the Eq. (6) will be created.

$$V_{i,j}=x_{i,j}+\Phi_{i,j}(x_{i,j}-x_{k,j}) \quad (6)$$

In Eq. (5) Φ factor represents a random value between [-1, 1], $j\in\{1, 2, \cdots, D\}$, $i\in\{1, 2, \cdots, SN\}$, and $x_{k,j}$ is a neighborhood for $x_{i,j}$ in population and $i\neq k$. Finally, with greedy selection between x_i and v_i based on fitness value of them, the one with more fitness value will be selected as a *i_th* place in population. Fitness value of each solution is calculated by the Eq. (7).

$$fit_i=\begin{cases} fit_i=\dfrac{1}{1+f(x)_i}\ iff(x)_i\geq 0 \\ 1+\left|f(x)_i\right|\ iff(x)_i<0 \end{cases} \quad (7)$$

In which, $f(x)_i$ is the amount of objective function for *i_th* solution. For each solution there is a counter that indicates the number of iterations that for them there is no improvement in this solution. Counter initial value is zero. If there was an alternative for the *i_th* solution the solution's counter becomes zero. Otherwise, a number is added to the counter. After all employed bees complete the search process; they share the information of the solutions and their position information with the onlooker bees on the dance area. This is equivalent to assessing a probability to each solution in memory. The probability of each solution is calculated according to the Eq. (8).

$$P_i = \frac{fit_i}{\sum_{i=1}^{SN} fit_i} \tag{8}$$

Onlooker bee based on probability of solutions, choose a solution and with using Eq. (6) searches for neighborhood (v_i) for this solution. According to the probability equation, the solution with higher probability has more chance to choose by onlooker bee.

Then with greedy selection based on fitness between x_i and v_i, the one which is better will be replaced as a *i_th* solution in population. Like employee bees phase, if the replacement is made for the solution, its counter value becomes zero, otherwise the counter is incremented. In the phase of scout bee, to avoid being trapped in local optima, the solution which it's counter value is greater than limit value, abandoned and employee bee that was responsible for it becomes scout bee. Then with random searching according to the Eq. (5), selects a new solution and in memory replaces it with removed one and the counter of the omitted one return to the zero. Usually the limit value is determined as a product of $SN \times D$. In each iteration of algorithm the best solution is kept (Kıran and Fındık, 2015; Karaboga, 2005; Kalwani *et al*., 1990). For more information about ABC, readers are referred to Karaboga and davis (2005) (Baghchesaraei, 2016).

2.3 Hybrid model of PNN-ABC

In this model, PNN have been used to model and predict brand share, but the difference between classic model and hybrid one is, in hybrid model it has been tried to find the best variance for classes. For this purpose, ABC has been applied.

First step: data entry

At this step, the required data of the problem is entered. It should be noted that the data must be suitable and free of any distorted data.

Second step: creation of a random population

First, a random population of 10 food positions (random possible solutions) is created for the standard deviation of each class (there are six classes in the present case study) within the range of .0001 to 2. These values are considered as the standard deviation of the PNN. In fact, in each repetition a new PNN is produced and the model is implemented. In each iteration, the number of classification errors is obtained (As it was mentioned earlier, bee colony algorithm parameters are determined using trial and error method and user skills. In the current case study, the number of 10 initial populations and the range of .0001 to 2 have been selected after several times running the model with various values because they result in the best answers).

Third step: Calculation of the fitness of each population member

Value of the fitness function should be calculated for every member of the population. Here, the number of misclassifications is calculated (Eq. (9)). Since this number should be minimized and the algorithm seeks a position with maximum fitness function, so the fitness function of each member of the population is obtained via Eq. (10):

$$fit_{\mathrm{i}} = \text{number of misclassification} \tag{9}$$

The population members should maximize the following relation:

$$fit_{\mathrm{i}} = \frac{1}{1 + fit_{\mathrm{i}}} \tag{10}$$

The maximum value of the above equation is considered as the superior answer. At this step, the initial σ is applied in PNN.

Fourth step: creation of a new position in the neighborhood of employed bees

Employed bees change one of the parameters to determine the new position. Employed bees provide the onlookers with information about the positions. Onlookers seek the points with the highest amount of nectar and they may not find a better answer. As a result, scouts look for a better answer. New answers are produced using the Eq. (11) with regard to the available initial population. Here, there are the number of j dimensions (equal to the classes) and i and k range from 1 to 10. New σ_s are calculated.

$$V_{ij} = X_{ij} + \Phi_{ij}(X_{ij} - X_{kj}) \tag{11}$$

Fifth step: calculation of the fitness function for the changed position

The value of fitness function for new answers is calculated.

Sixth step: greedy selection

Greedy selection is conducted between X_{ij} and V_{ij}. If the fitness function V_{ij} is better than X_{ij}, it will be selected and will replace X_{ij}. The fitness function V_{ij} should be compared with the 10 initial positions; otherwise, a new food source is created similar to the method mentioned in the previous step.

If all the observer bees (onlookers) have not been distributed, the algorithm will be transferred to the seventh step and if they have been distributed, the algorithm will be transferred to the tenth step.

Seventh step: Calculation of P_i values for X_i positions
P_i (Probability) is calculated according to the Eq. (8). This probability actually represents the possibility of a position is selected by the onlooker bees (In fact, P_i is the nectar of the food source in the main algorithm is the situation. The more the nectar, the greater the likelihood that the onlooker bees go to that position).

The number of 10 food sources has been considered in this case study.

Eighth step: X_i positioning with the highest P_i value
As it was mentioned above, the position with a higher P_i contains more nectar. In other words, this is the best position in real problems and is closer to the optimal solution.

Ninth step: change of the determined position of X_i
The onlooker creates a new position in the neighborhood of point P_i.

Tenth step: reaching the value of limit
All the above steps are repeated to achieve the best answer or ultimate iterations. If the value of limit was not achieved and the best answer was obtained: In this case, X_i would be selected as the best food position when the number of cycles has been finished. In addition, the algorithm will return to the fourth step when the number of cycles has not been finished.

If it reached the value of limit, an answer higher the limit (it equals 100 in the present case study) was repeated, and no improvement occurred; then, the algorithm progresses to the eleventh step (The parameter's value of limit has been obtained by the user based on trial and error method).

Eleventh step: creation of a new food position using Eq. (12)
In this step, a new answer is created by scouts according to the Eq. (12).

$$X_i^{j(\text{new})} = \min x_i^j + \Phi(\max x_i^j - min\, x_i^j) \tag{12}$$

Twelfth step: calculation of fitness function for new food positions
In this step, the values of fitness function for new positions are assessed and examined. The best position is kept in memory and the above steps are repeated until the optimal solution is achieved.

2.3.1 Schematic View of the Model

In Figure 2 below, the proposed model is presented in the form of a flowchart.

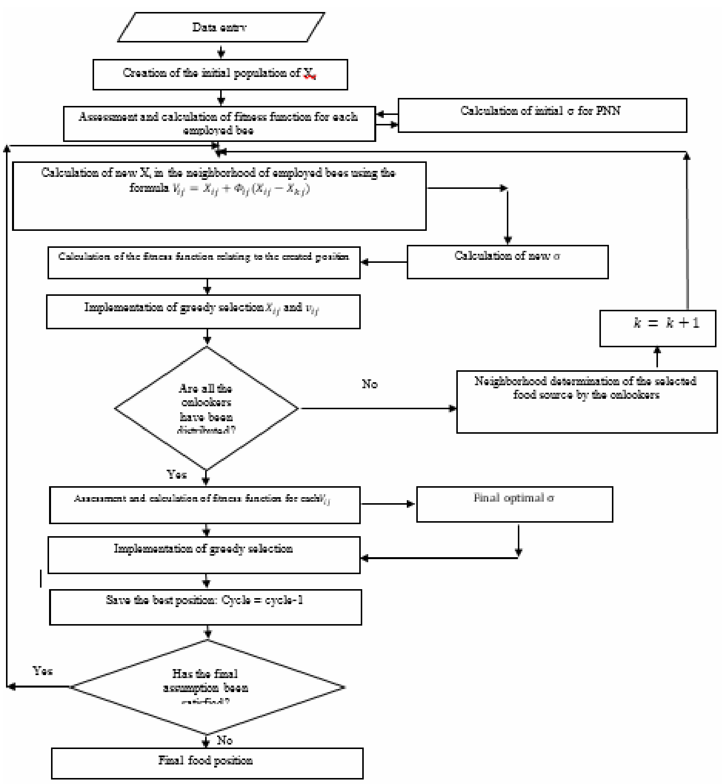

Figure 2. Flowchart of the hybrid model of PNN-ABC.

3. CASE STUDY OF PNN-ABC HYBRID MODEL

3.1 Data

In this section, we use the same data as those have been used by Kalwani *et al*. in 1990, Fish *et al*. in 2004, and we have used to model PNN-ABC for customer choice. These data have been used since they pertain to three various brands and two sizes; therefore, they suit the customer choice problem, which examines the customer choice among the available brands also customer model choices have been used for predict brand share. In addition, the size of the data is suitable for the implementation of neural networks because neural networks need a large volume of data for training and experimental phases. The data include coffee purchases by customers on an individual basis from four supermarkets in Kansas City over a period of 65 weeks. It is noteworthy that the data of 25 weeks were used to calculate the size and brand loyalty and the remaining 40 weeks are used for fitness/training and holdout set. Data for the remaining 40 weeks are divided into two categories, 1,564 data for training and 1,236 data for the holdout set have been considered. These data were cleaned by previous studies. The decision variables and output variable are listed in the table below. Since the retailers individually improve coffee brands and sizes (discount), brand-sizes have been modeled as choice options. For example, Floger trademark contains large-sized three-pound packages and small-sized one-pound packages in the dataset. Each of these compounds constitutes a different brand-size. The six compounds with the highest sale rates in the market (three brands, each containing two sizes) were maintained and the sum of these six compounds included 82.4% of the total number of purchases and 87.1% of market share. Each of the brand-sizes that were removed have taken up than 1% of the market share.

For an in-depth discussion of model's independent variables including their theoretical foundations, the reader is referred to Guadagni and Little (1983).

The data should be normal to implement the models. To this end, all the data should lie in the range [0, 1] so that all the data can be placed in the same range. Indeed, the effects of the data in a larger range may exceed the limit. For this purpose, the datum of each Eigen has been divided by the largest amount of data in that Eigen and the processing operation of the data is done.

3.2 Prediction of Brand Share

Since customer choice models are mostly used for the prediction of brand share and marketing policies, the production of new products, etc., the predictive power of each of the proposed models is compared in this section. To this end, mean absolute error has been used as a scale for the comparison of the models. Mean absolute error is calculated by obtaining the difference between the response predicted by the model and the real answer for each brand-size option.

Since the problem consists of 7 independent variables and 6 choice modes for each customer, there is the number of 42 input nodes. Agrawal *et al*.; and Fish *et al*. had obtained ANN with one hidden layer with 11 nodes through the optimization of the process. Furthermore, the number of hidden nodes is equal to the selected options, that is, 6. With the modeling and implementation of this network structure, the number of 626 errors was obtained after 1,500 iterations among the 1,236 test data.

In neural networks, the model accuracy and complexity increase and computing speed comes down in parallel with increasing the number of hidden layers. Thus, a neural network with two hidden layers was built so that accuracy would increase. In this way, the number of errors reached 539 in the 1,500-th iteration.

Since the problem contains 7 independent variables and 6 selection states for each customer, there will be the number of 42 input nodes for PNNs. The next layer in the PNN is the one in which all the 42 input nodes are connected to the 1,564 data of the problem and the 1,564

Table 1. Input and output variables for customer choice problem

Input variable	Description
$RP_{ik}(n)$	Original price of option i at the time of the n-th purchase by the k-th customer
$PROM_{ik}(n)$	Variable 0 and 1, which indicates the presence or absence of discounts in the price of option i at the time of the n-th purchase by the k-th customer
$PCUT_{ik}(n)$	The discounted price of option i at the time of the n-th purchase for the k-th customer
$PRV_{ik}(n)$	Variable 0 and 1, which indicates whether the previous purchase of the k-th customer has been followed by a discount on the option with the same brand i at the time n or not
$SPRV_{ik}(n)$	Variable 0 and 1, which indicates whether the two previous purchases of the k-th customer has been followed by a discount on the option with the same brand i at the time n or not
$BL_{ik}(n)$	Customer loyalty to the brand of the i-th choice at the time of the n-th purchase
$SL_{ik}(n)$	Customer loyalty to the size of the i-th choice at the time of the n-th purchase
Output variable	
$BC_{ik}(n)$	Selection of the i-th option by the k-th customer at the time of the n-th purchase (Karaboga, 2005)

data are divided into 6 classes (this problem has 6 output nodes). Then, the total units of each group exist and, at the end, the group with the highest aggregate is chosen.

The PNN network has 522 errors after modeling. The PNN has been used for the hybrid model because the number of errors is lower than that of feed-forward multilayer neural network with recursive back-propagation training and the computing speed is also higher.

In the next stage, the neural network containing 42 input nodes, 11 hidden nodes, and 6 output nodes is trained by the genetic algorithm to determine the pertaining weights. Fish *et al.* in 2,004 did so in their study. Their model was implemented; and 3,000 iterations and 440 errors were recorded where the presented answer is better than the neural network with back-propagation training.

Finally, PNN-ABC hybrid model is implemented for brand share predicting. Bee colony algorithm includes the following control parameters:

1) CS: It includes the employed bees (E_b) and onlooker bees (O_b)
2) The value of limit: It entails the number of attempts to stop the food source
3) The maximum cycle number (MCN)

Although ABC algorithm has three parameters that must be set, when the CS parameter is determined by the user for the first time, the value of limit is easily calculated. As a result, ABC algorithm comes out with two parameters that must be set: CS and MCN values.

Here, the number of CS has been considered equal to 20. Thus, an initial population equal to 10 is created (the initial population is equal to the number of the employed bees). In addition, MCN is considered equal to 1,458.

For the conduct of this experiment, the data are subdivided into three categories of training data, validation data, and holdout data.

Validation data have been used since they direct the algorithm towards better responses when implemented. To this end, 20% of the training data has been dedicated to the validation set, which is almost equal to 313 samples. The number of 391 errors occurred in iteration 1,458 after running the model.

In Table 2, mean absolute error values of different models after the number of iterations with the highest improvement are presented.

Table 2. Comparison of mean absolute error values of different models

Model	Number of iterations	Mean absolute error
ANN with one hidden layer	1,735	5.343
ANN with two hidden layers	1,500	4.672
PNN	1,600	3.894
ANN trained by GA	3,145	2.5502
PNN trained by ABC	1,458	2.3129

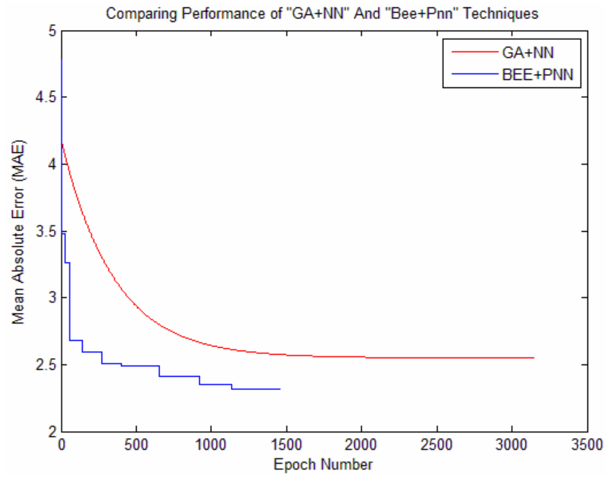

Figure 3. Comparison of PNN-RBC and ANN-GA performance.

As it is clear from the above table, PNN-ABC model has achieved a lower MAE with fewer iterations and this is indicative of the higher speed and accuracy of the model. Since PNN-ABC and ANN-GA models have the lowest MAE, the MAE change trend of PNN-ABC model has been compared with that of ANN-GA method. As it is shown in the Figure 3, PNN-ABC model has owned a lower MAE since the beginning of the model implementation and, thereby, it achieves proper MAE thro-ugh less iteration.

According to the results obtained from the calculation accuracy, the lower number of iterations, speed, and MAE; it can be concluded that PNN-ABC model has outperformed the other models in the modeling and predicting of consumer choice in the marketplace. For this reason, this model can be used to explain the marketing strategies, pricing strategies, etc. Moreover, company managers can be provided with this model. Therefore, in the following section, price sensitivity analysis has been used to show the power of this model in predicting customers' purchasing behavior.

3.3 Validation of PNN-ABC Model Using Price Sensitivity Analysis

According to the results obtained in the previous section, PNN-ABC model took advantage of higher accuracy and speed in predicting customer choice; therefore, in this section, price sensitivity analysis has been conducted to demonstrate the usability and applicability of PNN-ABC model. For this purpose, 25 cents, 50 cents, 75 cents, and a dollar has been added to and then reduced from each of the six brand-size options in terms of price, respectively. In this way, the application of the proposed model is examined in terms of marketing and pricing. In Table 3, the price sensitivity analysis results obtained for each brand-size option are presented. In each of these columns, the number of customers who opt for this option due to the price change is given.

Table 3. Price sensitivity analysis using PNN-ABC model

-$1	-75 cents	-50 cents	-25 cents	+$1	+75 cents	+50 cents	+25 cents	With the initial price	Brand-size
451	442	436	402	238	243	256	269	349	**1-1**
97	93	81	72	10	11	13	17	58	**1-2**
712	702	691	636	433	239	448	464	544	**2-1**
263	251	232	181	81	83	87	93	156	**2-2**
455	448	431	396	222	231	243	249	348	**3-1**
161	153	146	127	73	75	78	82	109	**3-2**

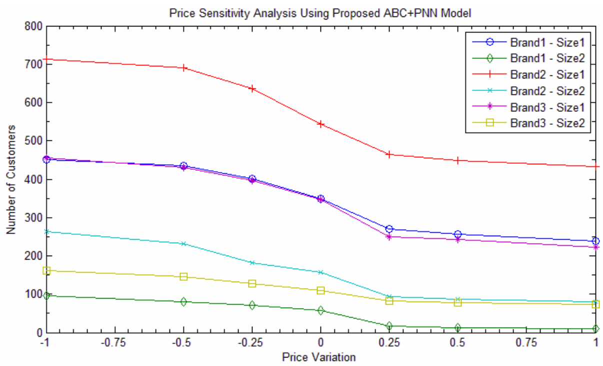

Figure 4. Results of price sensitivity analysis using PNN-ABC model.

The results are shown in Figure 4. As it is observed in the figure, PNN-ABC model displays a true predictability using price sensitivity analysis. In other words, there is an inverse relation between price and the number of customers, which means that the number of customers is reduced with rising prices and vice versa. Accordingly, it can be concluded that PNN-ABC model can be used as a strategic tool for the adoption of marketing policies and it can be offered to managers as the simulator of market reactions.

4. CONCLUSION

We had proposed a new approach by integration of Probabilistic Neural Network and Artificial Bee Colony for brand choice modeling. This hybrid model exhibited good results, higher accuracy and speed, compared with other models consist of Neural Network with one hidden layer, Neural Network with two hidden layer, Neural Network trained with Genetic Algorithm, and Probabilistic Neural Network. Since, usually brand choice models are used as a market response simulator, we used this hybrid model (PNN-ABC) to predict brand share of companies. In order to evaluate power of PNN-ABC in brand share prediction we have done a case study in real market. We used scanner-panel data which consist of f ground coffee purchases from four Kansas City supermarkets over a 65-week period. Results showed that PNN-ABC has higher accuracy in predicting brand share related to other approaches. Therefore PNN-ABC can be used as a suitable tool to deal with consumer behavior modeling and simulation.

REFERENCES

Agrawal, D. and Schorling, C. (1996), Market share forecasting: an empirical comparison of artificial neural networks and multinomial logit model, *J Retail*, **72**, 383-407.

Armano, G., Marchesi, M., and Murru, A. (2005), A hybrid genetic-neural architecture for stock indexes forecasting, *Information Sciences*, **170**(1), 3-33.

Baghchesaraei, A., Kaptan, M. V., and Baghchesaraei, O. R. (2015), Using Prefabrication Systems in Building Construction, *International Journal of Applied Engineering Research*, **10**(24), 44258-44262.

Baghchesaraei, O. R. and Baghchesaraei, A. (2014), Analytical survey of structural engineering and long-term resistive environmental elements in an Iranian magnificent palace, *International Journal of Civil and Structural Engineering*, **4**(3), 372-380.

Baghchesaraei, O. R., Lavasani, H. H., and Baghchesaraei, A. (2016), Behavior of Prefabricated Structures in Developed and Developing Countries, *Bulletin de la Société des Sciences de Liège*, **85**, 1229-1234.

Banerjee, A., Awasthy, D., and Gupta, V. (2005), A choice modelling approach to evaluate effective-ness of brand development initiatives, *International Journal of Management and Decision Making*, **6**(2), 180-198.

Barthwal, R. R. (2007), *Industrial Economics: an introductory text book*, New Age International.

Bates, J. M. and Granger, C. W. J. (1969), The combination of forecasts, *Journal of the Operational Research Society*, **20**(4), 451-468.

Bentz, Y. and Merunka, D. (2000), Neural networks and the multinomial logit for brand choice modelling: a hybrid approach, *Journal of Forecasting*, **19**(3), 177-200.

Borenstein, S. (1990), Airline mergers, airport dominance, and market power, *The American Economic Review*,

80(2), 400-404.

Buzzell, R. D., Gale, B. T., and Sultan, R. G. M. (1975), Market share-a key to profitability, *Harvard business review*, **53**(1), 97-106.

Chen, K.-Y. and Wang, C.-H. (2007), A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan, *Expert Systems with Applications*, **32**(1), 254-264.

Cheng, J.-H., Chen, H.-P., and Lin, Y.-M. (2010), A hybrid forecast marketing timing model based on probabilistic neural network, rough set and C4.5, *Expert systems with Applications*, **37**(3), 1814-1820.

Clemen, R. T. (1989), Combining forecasts: A review and annotated bibliography, *International journal of forecasting*, **5**(4), 559-583.

Fish, K. E., Johnson, J. D., Dorsey, R. E., and Blodgett, J. G. (2004), Using an artificial neural network trained with a genetic algorithm to model brand share, *Journal of Business research*, **57**(1), 79-85.

Gan, C., Limsombunchai, V., Clemes, M., and Weng, A. (2005), Consumer Choice Prediction: Artificial Neural Networks versus Logistic Models, *Journal of Social Sciences*, **1**(4), 211-219.

Gorunescu, F. (2006), Benchmarking probabilistic neural network algorithms, *In Proceedings of International Conference on Artificial Intelligence and Digital Communications*, 1-7.

Guadagni, J. D. C. and Little (1983), A Logit Model of Brand Choice Calibrated on Scanner Data, *Mark Sci*., **2**, 203-238.

Hansen, G. S. and Wernerfelt, B. (1989), Determinants of firm performance: The relative importance of economic and organizational factors, *Strategic management journal*, **10**(5), 399-411.

Hruschka, H. (2007), Using a heterogeneous multinomial probit model with a neural net extension to model brand choice, *Journal of Forecasting*, **26**(2), 113-127.

Hruschka, H. Fettes, W., Probst, M., and Mies, C. (2002), A flexible brand choice model based on neural net methodology A comparison to the linear utility multinomial logit model and its latent class extension, *OR Spectrum*, **24**(2), 127-43.

Hu, M. Y. (2003), Tsoukalas C. Explaining consumer choice through neural networks: The stacked generalization approach, *European Journal of Operational Research*, **146**(3), 650-660.

Hu, M. Y., Shanker, M., Zhang, G. P., and Hung, M. S. (2008), Modeling consumer situational choice of long distance communication with neural networks, *Decision Support Systems*, **44**(4), 899-908.

Kalwani, M. U., Yim, C. K., Rinne, H. J., and Sugita, Y. (1990), A price expectations model of customer brand choice, *Journal of Marketing Research*, 251-262.

Karaboga, D. (2005), An idea based on honey bee swarm for numerical optimization, Technical report-tr06, *Erciyesuniversity, engineering faculty, computer engineering department*, **200**.

Karaboga, D. and Akay, B. (2007), Artificial bee colony (ABC) algorithm on training artificial neural networks, In *IEEE 15th Signal Processing and Communications Applications*.

Karaboga, D. and Akay, B. (2009), A comparative study of artificial bee colony algorithm, *Applied Mathematics and Computation*, **214**(1), 108-132.

Karaboga, D. and Basturk, B. (2007), A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm, *Journal of global optimization*, **39**(3), 459-471.

Karaboga, D. and Basturk, B. (2007), A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm, *Journal of Global Optimization*, **39**(3), 459-471.

Karaboga, D., Gorkemli, B., Ozturk, C., and Karaboga, N. (2014), A comprehensive survey: artificial bee colony (ABC) algorithm and applications, *Artificial Intelligence Review*, **42**(1), 21-57.

Kaya, T., Aktas, E., Topçu, İ., and Ülengin, B. (2010), Modeling toothpaste brand choice: an empirical comparison of artificial neural networks and multinomial probit model, *International Journal of Computational Intelligence Systems*, **3**(5), 674-687.

Kazemi, S. M. R., Hadavandi, E., Mehmanpazir, F., and Nakhostin, M. M. (2013), A hybrid intelligent approach for modeling brand choice and constructing a market response simulator, *Knowledge-Based Systems*, **40**, 101-110.

Khashei, M., Hejazi, S. R., and Bijari, M. (2008), A new hybrid artificial neural networks and fuzzy regression model for time series forecasting, *Fuzzy Sets and Systems*, **159**(7), 769-786.

Kıran, M. S. and Fındık, O. (2015), A directed artificial bee colony algorithm, *Applied Soft Computing*, **26**, 454-462.

Kumar, A., Rao, V., and Soni, H. (1995), An empirical comparison of neural network and logistic regression models, *Marketing Letters*, **6**(4), 251-263.

Lee, W.-I., Shih, B.-Y., and Chung, Y.-S. (2008), The exploration of consumers' behavior in choosing hospital by the application of neural network, *Expert systems with applications*, **34**(2), 806-816.

Reid, M. J. (1968), Combining three estimates of gross domestic product, *Economica*, **35**, 431-444.

Specht, D. F. (1990), Probabilistic neural networks and the polynomial adaline as complementary techniques for classification. Neural Networks, *IEEE Transactions on*.

Specht, D. F. (1990), Probabilistic neural networks, *Neural networks*, **3**(1), 109-118.

van Wezel, M. and Potharst, R. (2007), Improved customer choice predictions using ensemble methods, *European Journal of Operational Research*, **181**(1), 436-452.

Vroomen, B., Franses, P. H., and van Nierop, E. (2004), Modeling consideration sets and brand choice using artificial neural networks, *European Journal of Operational Research*, **154**(1), 206-217.

Wasserman, P. D. (1993), Advanced methods in neural networks, *Chapter*, **3**, 35-55.

West, P. M., Brockett, P. L., and Golden, L. L. (1997), A comparative analysis of neural networks and statistical methods for predicting consumer choice, *Marketing Science*, **16**(4), 370-391.

Wierenga, B., van Bruggen, G. H., and Althuizen, N. A. P. (2008), Advances in marketing management support systems. Springer US.

Xiang, W.-L. and An, M.-Q. (2013), An efficient and robust artificial bee colony algorithm for numerical optimization, *Computers and Operations Research*, **40**(5), 1256-1265.