JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 25, 1605-1616 (2009)

A New Perspective for Neural Networks: Application to a Marketing Management Problem

JAESOO KIM AND HEEJUNE AHN+
Department of Computer Science and Engineering
+Department of Control and Instrumentation Engineering
Seoul National University of Technology
Seoul, 139-743 Korea
E-mail: heejune@snut.ac.kr

Over the last few years, connectionism or neural networks (NN) have successfully been applied to a wide range of areas and have demonstrated their capabilities in solving complex problems. Current indications show that these techniques are very important and rapidly developing areas of research and applications, particularly, in the area of data mining for knowledge discovery. One particular neural network model, the back-propagation (BP) algorithm, has performed very well in this regard and it is now accepted as a reliable method for data mining. However, these models have their shortcomings. The major difficulty lies in the fact that the relationships between specific variables and the neural network results are, at best, difficult to explain. This article presents an innovative but simple method for using NN to understand the pattern/outcome correlation to interpret a cause and effect relationship. A comparative analysis and experimental results are also presented to show the validity of the proposed scheme.

Keywords: neural networks, sensitivity analysis, CART, logistic regression, data mining

1. INTRODUCTION

1.1 Background

We all know that information has become a very important commodity. Every second hundreds of thousands of new records of information are generated. This information needs to be summarized and synthesized if it is to support effective decision-making. This involves the challenge of dealing with huge sets of data, dynamic data, incomplete or imprecise data, noisy data and missing attributes, and redundant or insignificant data. A successful approach to modeling non-linear relationships under these situations can be usage of artificial neural networks (ANN) or connectionism that can be trained with the set of available data [1, 8]. One particular neural network type, the back-propagation (BP) algorithm has performed very well in this regard and it is now accepted as a reliable method for data mining [2].

Neural networks share the advantages with the many other data mining tools. An advantage they have over classical models used to analyzed data, such as regression analysis, is that they can fit data where the relation between independent and dependent variables is nonlinear and where the specific form of the nonlinear relationship is unknown. Also, decision trees, a method of splitting data into homogeneous clusters with

Received October 19, 2007; revised April 9 & July 17, 2008; accepted August 28, 2008.
Communicated by Chin-Teng Lin.
+ Corresponding author.

similar expected values for the dependent variable, are often less effective when the predictor variables are continuous than when they are nominal (or categorical). Neural networks work well with both nominal and continuous variables. They do not require that the relationships between predictor and dependent variables be linear whether or not the variables are transformed. The neural network method is more robust and has better predictive accuracy than classical methods, such as discriminant and logistic analysis, in many data mining applications. As the focus of this paper is neural networks, the other data mining techniques will not be discussed further.

In spite of their advantages, neural networks with BP algorithm have their shortcomings. The major difficulty lies in the fact that the relationships between specific or causal variables and the neural network results are, at best, difficult to explain because of the complexity of the functions used in the neural network approximations. The output of a neural network is a predicted value and some goodness of fit statistics. However, the functional form of the relationship between predictor and target variables is not made explicit. So the nature of the strength of the relationship between the independent and dependent variables, *i.e.*, the importance of each variable, is usually not revealed. Validating unexplainable results can be a significant challenge. This means there must be something more general in their activity that led them to this result. Here we face the challenge of finding appropriate way to figure out these interrelationships to make use of them in the future without requiring some additional knowledge about the character of the task.

Basically, the aim of this paper is to show that the neural network modeling may offer significant advantages over the commonly used estimation procedures that can summarise the large amount of collected data into relevant, concrete and effective action recommendations for decision makers. In order to meet the growing demand of decision makers, we have to focus on the systems that find adequate explanation models.

In this study, we also tested the comparative abilities of a neural network model, logistic regression, and classification and regression trees (CART) at capturing interrelationship between the independent variables and the dependent variables. And this paper describes how a model of factors influencing consumer behavior, from which initial measures can be used, can be produced using a neural network based on consumer survey data.

1.2 Cultural Orientation and Consumer Behavior

The culturally based norms (appropriate behavior in a situation) and values (desirable behavior across situations) would lead to differences in consumer behavior across cultures. These values and norms are passed on from the community to an individual as he or she is socialized within the community. Consumers learn values and norms about the acquisition, consumption and disposal of products through socialization in their communities. Thus cultural values and norms become a primary explanation of similarities in behavior of individuals within the community, and differences in the behavior of individuals across communities [4].

Especially, Chinese consumer behavior is essentially different because of its unique cultural, social and economic roots [13-15]. The behavior of Chinese consumers has even been distinguished from that of consumers in other Asian countries [5]. Sun & Collins [13] studied consumers' attitudes towards imported fruit in Guangzhou, finding that fruit attributes in relation to symbolic and hedonic values were of primary importance in the decision to purchase.

Imported fruits have been widely available in China since about 1993. They are retailed throughout the year in every major city in the country. Though imported fruit is more expensive than locally produced fruit, there are still many willing buyers and the imported fruit business has experienced burgeoning demand and high profits.

For this study, survey data were collected through structured intercept interviews with consumers at point of sale immediately after they had purchased imported fruit. Results will help to broaden our understanding of Chinese consumer behavior and provide valuable information when formulating marketing strategies.

The rest of the article is organized as follows. Section 2 reviews the studies on neural networks. In section 3, we introduce a sensitivity measure that assesses the relative importance of the input factors used by the network to arrive at its targets and review the existing relative importance measure for neural network input elements. In section 4, we apply the method discussed in section 3 to a marketing management problem. Then, section 5 compares and contrasts how neural networks and classical modeling techniques deal with the specific modeling challenges and how the output of neural networks can be used to better understand the relationship in the data through sensitivity analysis. Subsequently, we examine the results of our studies and test how the neural network model performs in practice using the real-world data set. Finally, we draw conclusions in section 6.

2. NEURAL NETWORK AND ITS UNDERSTANDING THE OUTPUT

Neural networks are based on an early model of human brain function. Although it is described as a network, a neural network is nothing more than a mathematical function that computes an output based on a set of input variables. The network paradigm makes it easy to decompose the larger function to a set of related subfunctions, and it enables a variety of learning algorithms that can estimate the parameters of the subfunctions.

There are many different types of neural networks. A feedforward neural network with one hidden layer considered in this paper is known as a multilayer perceptron (MLP), which is one of the most popular kinds of neural networks and uses supervised learning. As a result, its effectiveness has been established and software for applying it is widely available. It has also been proved that a network with only one hidden layer is enough to approximate any continuous function given there are enough nodes in the hidden layer [3, 9]. The hidden layers are used to model the nonlinearities in the relationship between inputs and output [11]. Therefore neural networks might represent a viable alternative to multivariate statistical methods.

Although neural networks can be applied to a number of data mining problems, including classification, regression, and clustering, the complexity, combined with the nondescriptive nature of neural network models, often discourages all but the most scientists and researchers from employing the data mining technique.

Neural networks are trained by adjusting weights by some automatic learning algorithms so that the result of stability approximates the desired outcomes for the provided inputs. The output from neural networks varies greatly. Other common outputs are accuracy measures such as confusion matrix, R^2, and so forth for validating the model. The output from a neural network is purely predictive. Unfortunately, none of these aids the user in understanding the model or the underlying data relationships. Because there is no descriptive component to a neural network model, a neural network's choices are hard to understand, and this often discourages its use. In fact, this technique is often referred to as a black-box technology.

Because of the more complicated functions involved in neural network analysis, interpretation of the variables is more challenging. One approach is to examine the weight connecting the input variables to the hidden layer. Those which are closest to zero are least important. A variable is deemed unimportant only if all of these connections are near zero. This procedure is typically used to eliminate variables from a model, not to quantify their impact on the outcome. Due to the homogeneous structure of neural network, it is hard to extract structured knowledge from either the weights or the configuration of the neural network in question. It should be emphasized that the weights in a neural network with hard-limiter as its activation function do have physical meaning [12]. The weights of a given node represent the coefficients of the hyperplane or discriminant function that partitions the input space into two regions with different output values. However, this interpretation of weights gets weaker and weaker if the net's activation function is either sigmoid or hyperbolic tangent functions and the given dependent variables are continuous instead of binary [10]. Therefore the weights are relatively uninformative for determining the influence of the variables on the fitted values.

Another approach to assessing the predictor variables' importance is to compute a sensitivity analysis for each variable. The sensitivity is a measure of how much the predicted value's error increases when the variables are excluded from the model one at a time. Through the sensitivity analysis, it is possible to generate an estimate of the general level of influence exhibited by each parameter from an analysis of the network weights in a systematic manner. This can be used to rank each variable's importance. In the following section, a method for calculating output sensitivities to inputs' variations from a trained neural network is discussed in some detail.

3. SENSITIVITY ANALYSIS

In general, one of the key factors that affect the success of process modeling is the ability to extract information about the model structure and the relationships between its inputs and outputs from the trained network. Such information is essential for model validation and for process optimization, control and safety assessments. Moreover, in some cases where the original process is not well understood, this information can be employed as a basis for the analysis of the process and in determining the most significant factors that affect it.

For multilayer feedforward networks with n input nodes, one hidden layer with h nodes and k output nodes the relative importance (RI) of the ith component of the input vector can be estimated as follows:

$$\boldsymbol{RI}_{ik} = \frac{\sum_{j=1}^{h}\left|\frac{\boldsymbol{w}_{ji}\boldsymbol{w}_{kj}}{\sum_{i=0}^{n}\left|\boldsymbol{w}_{ji}\right|}\right|}{\sum_{j=0}^{h}\left|\boldsymbol{w}_{ji}\right|}, \tag{1}$$

where w_{ji} is the weight from the ith input node to the jth hidden node and w_{kj} is the weight from the jth hidden node to the kth output node. Biases are given the subscript 0.

Hence, the *RI* measure incorporates certain rates of change of the strengths of signals as they flow through the network. For example, w_{ji} is the partial derivatives of the inputs to the hidden layer with respect to the inputs to the network. Similarly, w_{kj} is the partial derivatives of the inputs to the output layer with respect to the outputs of the hidden layer. So this *RI* measure is simply compounded weighted averages and is independent of the activation function, therefore it is applicable to networks trained on a range of activation functions, which are monotonically increasing.

Eq. (1) includes a component to normalize for the effect of extreme weights connecting input and hidden nodes. This additional component is also included in a closely related formula given by Garson [7]:

$$\boldsymbol{RI}_{ik} = \frac{\sum_{j=1}^{h}\frac{\left|\boldsymbol{w}_{ji}\right|\left|\boldsymbol{w}_{kj}\right|}{\sum_{i=1}^{n}\left|\boldsymbol{w}_{ji}\right|}}{\sum_{i=1}^{n}\sum_{j=1}^{h}\frac{\left|\boldsymbol{w}_{ji}\right|\left|\boldsymbol{w}_{kj}\right|}{\sum_{i=1}^{n}\left|\boldsymbol{w}_{ji}\right|\left|\boldsymbol{w}_{ji}\right|}}. \tag{2}$$

Thus, for each j of h hidden nodes, sum the product formed by multiplying the input-to-hidden connection weight of the input node i of variable for hidden node j, times the connection weight of the output node k for hidden node j, then divide by the sum of such quantities for all variables. The result is the percentage of all output weights attributable to the given independent variable, excepting bias weights arising from the back-propagation algorithm.

However, the related method proposed by Garson does not include the effect of the bias, which could result in a significant omission. Garson’s measure places more emphasis on the connection strengths from the hidden layer (w_{ji}) to the output layer (w_{kj}), but it does not measure the direction of influence (positive or negative). That is, during the summation process, positive and negative weights can cancel their contribution or influence, which leads to inconsistent results. Including the bias influence allows all the influences to be considered in the context of the complete network. For instance, it is possible, although unlikely, that the output of a network is based purely on the bias, and the input signal has no significant effect. Using Garson's approach, the input parameters could be assigned influence to various degrees since the overwhelming bias effect is ignored.

The *RI* measure given in this paper would illustrate the minimal (zero) influence of the inputs and the large effect of the bias. In this way, it is possible that the denominator in Eq. (1) will reduce to zero for non-zero weights. That is, the denominators will only be zero if all the weights are zero, for instance all weights from the hidden layer to the output layer are zero resulting in a network which simply outputs a single value determined by the activation function for any input signal or for all weights from the input node under consideration to the hidden layer, including the bias, to be zero in which case the input parameter will have no effect. Also, note that the numerator can become zero under the same conditions, *i.e.*, non-zero weights. In this case the strength of the input would be zero. In the following section, it will be shown how this method can be applied to a real world problem.

4. APPLICATION EXAMPLE

4.1 Data Collection and Statistical Information

For this study, a supermarket in China was chosen and the survey conducted was mall-intercept personal interviews. The shop in its fruit section was deliberately divided commodities into two subsections, domestic fruit and imported fruit, and also marked in both sides with clear sign. It was a heavily trafficked store, and its management gave approval to promote the survey as being on behalf of the company.

The purpose of the survey was explained to interviewees as being for the improvement of service so that customers' needs could be better understood and met. This was in fact true because the company wanted to use the results. The rationale for this approach was to ensure that the interviewee's personal interest was directly associated with the quality of their answers to the questions. Surveys were administered so as to avoid public holidays and to achieve a spread across weekdays. The survey involved 520 personal interviews in Guangzhou and a total of 495 useable responses was recorded for the study.

Respondents were asked about both their *beliefs* and *evaluation* of the imported fruit, including their intention of purchasing imported fruit on each statement. 11 questions relating to consumers' attitudes and perceptions that might be motivated their buying intention to imported fruit were designed.

The questions that consumers were asked in relation to these 11 attributes were framed into statements according to Fishbein's theory [6]. Fishbein's proposition is that people form attitudes towards a product attribute on the basis of their belief about that attribute (comprised of perceptions and knowledge) and their positive or negative feelings towards that attribute (comprised of their evaluation of that belief). According to Fishbein, a consumer's overall attitude toward imported fruit products would be represented by the sum of the products of their beliefs about each attribute and their evaluation of those beliefs.

Among them, seven were the objective characteristics relating to attributes and perceptions of imported products, such as *appearance*, *packing*, *pollution*, *taste good*, *taste different*, *freshness* and *price*. Four were the subjective attributes towards symbolic means of purchasing imported fruit products: *achievement*; *wealthy*; *personality* and *social statues*. Besides assessing consumers' attitudes and perceptions of imported fruit a behavioral response measure of consumer intention was elicited.

Table 1 describes some features of the independent variables and the dependent variables used in this work. Note that there is no missing for the variables. Some descriptive statistics for the data set of the consumers' attitudes and perceptions towards im-

Table 1. Description of attributes: the independent variables ($x_1 \sim x_{11}$) and the dependent variable (y).

Attributes	Variable Type
Appearance of fruit (x_1)	continuous
Packing status (x_2)	continuous
Less pollution (x_3)	continuous
Taste good (x_4)	continuous
Taste different (x_5)	continuous
Freshness (x_6)	continuous
Person's achievement (x_7)	continuous
Wealthy (x_8)	continuous
Personality (x_9)	continuous
Social status (x_{10})	continuous
Attitude to price (x_{11})	categorical
Buying intention (y)	categorical

Table 2. Descriptive statistics: each value is the product of beliefs by evaluation.

Attributes	Mean	Std. Dev.	Skewness
Appearance of fruit	5.54	4.39	1.76
Packing status	5.70	4.78	1.73
Less pollution	7.60	5.57	1.00
Taste good	7.72	5.91	1.08
Taste different	6.80	5.53	1.51
Freshness	6.42	5.45	1.26
Person's achievement	13.51	7.78	0.18
Wealthy	13.64	7.77	0.17
Personality	12.37	7.98	0.33
Social status	13.91	8.06	0.08
Attitude to price	1.71	0.46	– 0.91

ported fruit are also illustrated in Table 2. The data in this example have skewness values of ranging from 0 to 2, which are considered acceptable for this task so that approximate normality is attained after the data is logged.

4.2 Methods

The data set consists of 495 respondents on consumer survey in an imported fruit market. The cross-validation procedures used in the neural network simulation were applied to the questionnaire data to prevent overfitting; that is 30 percent of the sample was used to train the network, 20 percent was used to determine a stopping point for training, and the remaining 50 percent was used for hold-out-sample testing of the predictive accuracy. For the logistic regression and the CART modeling, the data was separated into a training set of 248 customers and a test set of 247 customers. Variables used are described in Table 1.

The logistic regression coefficients correspond to "*B*" coefficients in the logistic regression equation indicate the amount of change expected in the *log odds* when there is a one unit change in the predictor variable with all of the other variables in the model held constant. A coefficient close to 0 suggests that there is no change due to the predictor variable. There is a relationship between the logistic coefficients and the odds ratios, *odds ratio* = *Exp*(*B*). These coefficients are used to compare the relative importance of the independent variables in this work as it can be seen in Table 3.

Table 3. Coefficients of the relative importance to the various imported fruit characteristics.

Attr	Logistic *Exp*(*B*) (rank)	NN *RI* (rank)	CART *Scores* (%) (rank)
x_1	1.062 (3)	0.015 (11)	3.71 (10)
x_2	0.928 (10)	0.072 (7)	0.00 (11)
x_3	1.060 (4)	0.266 (2)	27.55 (3)
x_4	1.107 (2)	0.227 (3)	100.0 (1)
x_5	0.996 (8)	0.066 (8)	25.22 (4)
x_6	0.926 (11)	0.028 (10)	8.21 (9)
x_7	1.030 (5)	0.111 (4)	22.35 (6)
x_8	1.019 (6)	0.093 (5)	22.14 (7)
x_9	0.991 (9)	0.042 (9)	11.34 (8)
x_{10}	1.005 (7)	0.087 (6)	24.50 (5)
x_{11}	2.073 (1)	0.269 (1)	57.38 (2)

For building a neural network model, we only considered a feedforward with a single hidden layer architecture, as they can approximate any continuous function and training algorithm was the *Lavenberg-Marquardt* algorithm. The size of hidden nodes needs to be only a relatively small fraction of the input layer. In this study, one empirical guideline is to determine the number of hidden nodes as twice the square root of the sum of input and output nodes, design multiple networks by varying the initial weights, and use the validation set to choose the best network. If the network fails to converge to a solution, it may be that more hidden nodes are required. If it does converge, we may try fewer hidden nodes. Application of the fitted model to the test data indicated that a 4 node neural network provided the best model (*i.e.*, 11-4-1). The performance measures of each neural network model such as MSE, and classification accuracy (%) are the average of 5 trials.

To calculate a variable importance score, CART looks at the improvement measure attributable to each variable in its role as a surrogate to the primary split. The values of these improvements are summed over each node and totaled, and are scaled relative to the best performing variable. In such ways, CART automatically produces the variable importance ranking (scores) based on the contribution predictors make to the construction of the tree. The variable importance rankings or predictor rankings (%) are strictly relative to a specific tree; change the tree and we might get very different rankings.

5. COMPARATIVE ANALYSIS AND RESULTS

In this example, we attempt to assess the relative importance of input or causal variables by examining a trained neural network model using the real-world data. A neural network with 4 nodes in the hidden layer was run on a training and validation set. Each of the three models was tested on consumer survey data and used to rank variables in importance.

Table 3 displays the results of the sensitivity test for each of the variable and shows the coefficients of the models developed with the logistic regression method, ranking scores (%) with the CART method and the neural network method using Eq. (1). The table shows that x_{11} (price) is the most important input factor, followed closely by x_3 (less pollution); x_4 (taste good) is of substantial but somewhat lesser importance. On the basis of this measure, we could also correctly infer that in the case of subjective characteristics, x_7 (achievement) is the most likely direct cause of *buying behavior* and that x_8 (wealthy) are slightly less important but still substantial causes of the consumer's *buying behavior*.

The odds ratio (or β weights or $Exp(B)$) in regression is interpreted as the ratio of the relative importance of the causal or input variables in the model. As indicated by the $Exp(B)$ for the *odds ratio* in the table, x_{11}, x_4, x_1, x_3 and x_7 have the highest importance in affecting the choice of the consumer's buying strategy. For this example, the *odds ratio* of x_{11} was 2.073, 1.107 was x_4, and 1.062 was x_1. On the basis of regression, we would correctly infer that x_{11} was the most important and likely direct cause of *buying intention* and that x_4, x_1, x_3 and x_7 were equal but lesser causes. This similarly matches with the results from the neural network analysis.

In the CART model case, the scores (%) reflect the contribution each variable makes in classifying or predicting the dependent variable, with the contribution stemming from both the variable's role as a primary splitter and its role as a surrogate to any of the primary splitters. In this example, x_4, the variable used to split the root node, is ranked as most important. The variable, x_2, received a zero score, indicating that this variable did not play any role in the analysis as either as primary splitters or surrogates. x_{11} (price) also is an important splitting variable and has the highest scores, followed by x_3, x_5, x_{10} and x_7 as illustrated in Table 3.

For all three models' measures, x_{11} has the highest contribution. For both our measure and in logistic regression's measure, x_3 and x_4 were second, respectively, whereas in CART, x_{11} was second, Thus, the logistic and neural network methods identify x_{11} as the most important variable, whereas the CART model identifies x_4 as the most important variable.

According to the sensitivities, x_3, x_4 and x_{11} are the most important variables and x_1 is the rather least important. This contrasts with the importance rankings of x_1 in the logistic analysis, where x_1 was a more important variable than others. Note that these are the sensitivities for the particular models. A different initial starting point for the neural network or a different number of hidden nodes could result in a model with different sensitivities and the rankings can be quite sensitive to random fluctuations in the data. The importance rankings in CART need to be understood as being relative to a particular tree and the rankings are strictly relative to a given tree structure. Overall, it shows that the consumer's objective characteristics were more important than the subjective characteristics in effecting the consumers' purchasing behavior. The low sensitivities were probably a result of the high correlations of the variables with each other.

These findings have obvious managerial implications since the consumers' perceptions of the various imported fruit characteristics can be influenced by managerial actions. For some variables (*e.g.*, *packing status* and *freshness*), the financial costs needed to change consumers' perceptions might be large, but these results show that it may have little effect on consumer's purchasing intention and should therefore not be a priority item for managerial action. Conversely, since the ultimate purchasing behavior is more strongly influenced by other variables, such as *price*/*taste good*/*less pollution*, managerial actions affecting the consumers' perceptions and attitudes of the imported fruit selection can better influence consumer's buying expectations in the store.

Finally, the classification results for all three models are presented in Table 4. The cross-validation procedure described earlier was used for all three models. The results in Table 4 demonstrate that once again the neural network model exhibits a superior ability to learn the patterns corresponding to consumer choice (buying intention). Consistent with the simulation results, the neural networks demonstrate significantly better hold-out-sample predictive accuracy that of the other models.

Table 4. Summary of classifications of the consumers' attitudes and perceptions by logistic regression, neural network and CART.

Analysis Method	Training	Hold-Out
Logistic Regression	67.7%	63.56%
Neural Network	89.59%	65.18%
CART	82.66%	61.13%

6. CONCLUSIONS

In this paper, we have developed a method for determining the relative importance of each input or causal variable of a neural network on the target. We have then applied this method to a neural network model of an empirical examination of consumer's behavior and showed that the neural network models can be used to improve the predictions to an important business management problem as well as understand the relative causal importance and order of the input variables.

Neural networks seem to have an advantage over linear models when they are applied to complex nonlinear data and may outperform classical models in certain situations, but interpreting the result is difficult because the nature of the relationship between dependent and target variables is not usually revealed. Neural networks also have no problem with trigonometric or logarithmic relationships, but either of these could be a real problem for the other techniques. This is an advantage neural networks share with other data mining tools not discussed in detail in this paper.

A method for interpreting the results of neural networks is presented here and incorporating such method into neural network models would help address the limitation. Our *RI* measure does provide a reasonable method of using neural network for modeling as well as for classification or prediction, and stands in sharp contrast to misleading views of neural networks as black-boxes whose iterative processes are beyond human compre-hension, even if the predictions are good.

REFERENCES

1. C. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, New York, 1995.
2. K. J. Cios, W. Pedrycz, and R. W. Swiniarski, *Data Mining Methods for Knowledge Discovery*, Kluwer Academic Publishers, Dordrecht, 1998.
3. G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control*, *Signals and Systems*, Vol. 2, 1989, pp. 303-314.
4. G. J. Tellis and D. S. Ackerman, "Can culture affect prices? A cross-cultural study of shopping and retail prices," *Journal of Retailing*, Vol. 77, 2001, pp. 57-82.
5. J. X. Fan and J. J. Xiao, "Consumer decision-making styles of young-adult Chinese," *Journal of Consumer Affaires*, Vol. 32, 1998, pp. 275-289.
6. M. Fishbein, *The Relationship Between Beliefs*, *Attitudes*, *and Behavior*, *Cognitive Consistency*, S. Feldman, ed., Academic, New York, 1966, pp. 199-223.
7. G. D. Garson, "Interpreting neural-network connection weights," *AI Expert*, 1991, pp. 47-51.
8. S. Haykin, *Neural Networks: A Comprehensive Foundation*, MacMillian and IEEE Computer Society, New York, 1994.
9. K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, Vol. 2, 1989, pp. 359-366.
10. M. Ishikawa, "Structural learning with forgetting," *Neural Networks*, Vol. 9, 1996, pp. 509-521.
11. C. Klimasauskas, "Neural networks: An engineering perspective," *IEEE Communication Magazine*, Vol. 30, 1992, pp. 50-53.
12. R. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Magazine*, Vol. 4, 1987, pp. 4-22.
13. X. Sun and R. Collins, "Attitudes and consumption values of consumers of imported fruit in Guangzhou, China," *International Journal of Consumer Studies*, Vol. 26, 2002, pp. 34-43.
14. X. Sun and R. Collins, "A comparison of attitudes among purchasers of imported fruit in Guangzhou and Urumqi, China," *Food Quality and Preference*, Vol. 15, 2004, pp. 229-237.
15. O. H. M. Yau, *Consumer Behavior in China*: *Customer Satisfaction and Cultural Values*, Routledge, London, 1994.

Jaesoo Kim (金在洙) is an Associate Professor at Department of Computer Science and Engineering, Seoul National University of Technology. He received his M.S. and Ph.D. degrees in Computer Science and Information Science from Monash University, Australia, University of Otago, New Zealand in 1992 and 1999, respectively, and his B.S. degree in Computer Science from Seoul National University of Technology, South Korea in 1988. Before joining Seoul National University of Technology as a Professor, he worked as an Assistant Professor at Zayed University, UAE from 2002 to 2003, and University of Queen-

sland, Australia from 2000 to 2001. Before his Ph.D., he worked as a research engineer at Otago University, New Zealand from 1997 to 1999, University of Auckland, New Zealand from 1993 to 1996, and Monash University, Australia from 1989 to 1992. His interests include artificial intelligence, computing intelligence, web based data mining, and intelligent business decision system.

Heejune Ahn (安熙準) received his Ph.D., M.S., and B.S. degrees in Electrical Engineering from KAIST (Korea Advanced Institute of Technology), Daejeon, the Republic of Korea, in 1999, 1995 and 1993, respectively. He is an Assistant Professor of the Department of Control and Instrumentation Engineering at Seoul National University of Technology, Seoul, the Republic of Korea. He worked as a visiting researcher at Telecommunication Lab. of Erlangen-Nuremberg University, Germany, from July 1999 to February 2002. He has been a GSM/GPRS/UMTS wireless mobile protocol software engineer at Next Generation Handset Lab., LG Electronics Inc., Korea, from February 2000 to September 2002. From September 2002 to December 2003 he worked as a software architect and programmer of J2EE web server system at Tmax Soft Inc. His research interests include multimedia communications, protocol development, network system performance analysis, and real-time embedded systems.