World Academy of Science, Engineering and Technology
International Journal of Economics and Management Engineering
Vol:8, No:9, 2014

Integrating Artificial Neural Network and Taguchi Method on Constructing the Real Estate Appraisal Model

Mu-Yen Chen, Min-Hsuan Fan, Chia-Chen Chen, Siang-Yu Jhong

***Abstract*—**In recent years, real estate prediction or valuation has been a topic of discussion in many developed countries. Improper hype created by investors leads to fluctuating prices of real estate, affecting many consumers to purchase their own homes. Therefore, scholars from various countries have conducted research in real estate valuation and prediction. With the back-propagation neural network that has been popular in recent years and the orthogonal array in the Taguchi method, this study aimed to find the optimal parameter combination at different levels of orthogonal array after the system presented different parameter combinations, so that the artificial neural network obtained the most accurate results. The experimental results also demonstrated that the method presented in the study had a better result than traditional machine learning. Finally, it also showed that the model proposed in this study had the optimal predictive effect, and could significantly reduce the cost of time in simulation operation. The best predictive results could be found with a fewer number of experiments more efficiently. Thus users could predict a real estate transaction price that is not far from the current actual prices.

***Keywords*—**Artificial Neural Network, Taguchi Method, Real Estate Valuation Model.

I. Introduction

WITH the continuous decrease of usable lands and increase of population on the earth, fewer natural resources can be evenly distributed to everyone. From the perspective of the law of supply and demand, the price of land becomes high relatively. With the rapid growth of the capital of each country, convenient public facilities continue to develop, attracting large crowds, resulting in less volume rate of urban real estate. Real estate prices soar in recent years, and more and more people see real estate as an investment target, causing turmoil in people's livelihood. Governments also actively intervene to control fluctuations in real estate. Residential demand is the most basic living condition for people. If the basic need cannot be taken care of, it is impossible to pursue steady development and high growth.

The application of artificial neural network and Taguchi method is no longer seen in few cases. In the past, many scholars also explored related issues. When artificial neural network is used in the training model, four main parameter settings have to be adjusted, namely, neurons, learning rate, momentum and epochs. When there is a suitable parameter combination, a more suitable predictive model will be established. There are a lot of ways to optimize parameters. This study used the Taguchi method neural network parameter optimization and the characteristics of the orthogonal array to set three levels for the four parameters. The L_9 orthogonal array was also used to calculate the S/N ratio (Signal to Noise Ratio) for every experiment to obtain the optimal parameter setting. By using the recommended parameter combination, a better measurement model can be established by the artificial neural network [1]-[6]

M.Y. Chen is with the Department of Information Management, National Taichung University of Science and Technology, Taichung 404, Taiwan (phone: 886-4-22196310; fax: 886-4-22196311; e-mail: mychen@nutc.edu.tw).

M.H. Fan and S. Y. Jhong is with the Department of Information Management, National Taichung University of Science and Technology, Taichung 404, Taiwan (e-mail: mfan@nutc.edu.tw, s1801b115@gms.nutc.edu.tw, respectively).

C.C. Chen is with the Department of Management Information Systems, National Chung Hsing University, Taichung 402, Taiwan (e-mail: emily@nchu.edu.tw).

II. Literature Reviews

A. Real Estate Introduction

In 1990, Professor Chin-Oh Chang presented the following nine features for real estate: immobility, local outdoor production, long-term use, non-standardized "commodity," huge amount of investment, regional demands, large transaction amount, low number of purchase, high sensitivity and weak adjustment of market supply and demand. From the above nine features, it is known that predicting real estate transaction prices is a very complex issue [7].

B. Real Estate Valuation Method

Local and foreign scholars have tried every possible way to use technology to assist real estate appraisal. The International Association of Assessing Office (IAAO) in 1979 officially defined mass appraisal as the use of standardized procedures and statistical tests on real estate data to conduct a systematic valuation as of a given date. More and more scholars also began to explore this area. In their book "Computer Assisted Mass Appraisal" McCluskey and Adair, scholars in Northern Ireland, mentioned that lands and buildings had different conditions in different countries[8]. They used a large number of data to assess the value of real estate in various countries as shown in Table I.

TABLE I
LARGE NUMBER OF DATA TO ASSESS THE VALUE OF REAL ESTATE IN VARIOUS COUNTRIES

Year	Nation	Subject matter	Valuation method	Conclusion
1968	United States	Residence	Multiple regression analysis	To assist government revenue
1974	Canada	Farmland	A valuation committee develops and establishes evaluation system	To assess the value of agricultural land
1979	New Zealand	Residence	Computer-aided system Multiple regression analysis technique	To assess residential real estate
1980	United States	Residential complex	Computer-assisted mass valuation system	To assist in the assessment of tax
1983	Canada	Farmland	Use of agricultural land revenue for analysis	To establish the newest agricultural database
1990	Sweden	Farmland	Use of geographic information system	Different valuation methodsfor different agricultural lands (arable land, pastures, forests)

C. Artificial Neural Network, ANN

Artificial Neural Network (ANN), originated in 1950, was used by scientists in half a century ago to mimic biological neural network structure, function and operation principle. In 1957, Rosenblatt began to develop artificial neural network models and a neuron model called "Perceptron" was established [9]. At the early stage, perceptron was applied in classifier, such as in sample identification or theoretical research. The main principle of operation is based on computer systems that mimic human thinking and behavior. Through learning, reasoning and judgment, it is possible for the system to analyze problems and remember ways of doing things. Scientists broke down this series of processes, analyzed every small section via computer programs and then applied it to the process of how humans deal with things. These actions then were modularized or formulized, so that the computer can systematically or structurally deal with more complex issues. Table II describes the development of artificial neural networks.

TABLE II
DEVELOPMENT OF ARTIFICIAL NEURAL NETWORKS

Year	Development process	Author
1943	Psychologists and mathematicians first used neurons as operation principle to propose a mathematical model.	McCulloch & Pitts [10]
1949	Hebbian learning rule was proposed. If two neurons respond at the same time, the link between neurons will be enhanced. Hebbian learning rule established a learning mechanism for artificial neural.	Heeb [11]
1957	The Perceptron model was established, which first developed the concept of artificial neural network.	Rosenblatt [9]
1969	A book entitled "Perceptron" was established, exploring the dilemma why Perceptron was unable to break through regional optimal solution.	Minsky & Papert [12]
1976	The adaptive resonance theory network (ART) was proposed, quite famous in unsupervised network model.	Grossberg [13]
1982	Self-Organizing Map (SOM) was proposed. The variables that are input can find the corresponding logical network in the absence of rules.	Kohonen [14]
1982	John. Höpfel invented Hopfield neural networks (HNN).	Hopfield [15]
1986	The back-propagation neural network (Back-propagation) was invented, solving the logic operation problems, and leading the artificial neural network to another peak.	Rumelhart, Hinton, & Williams [16]

D. Taguchi Method

Taguchi method was proposed in the 1950s by Dr. Taguchi in Japan [17]. With a simple orthogonal experimental design and a clear analysis of variance, the main aim is to use a small amount of experimental data for systematic analysis. The results were significant and the product quality could be improved. Many Japanese companies at that time adopted Taguchi method to improve product manufacturing process, called quality engineering. The method received the highest honor on quality management in the world in 1962 - Deming Application Prize. In 1951,1953, and 1984, it won Deming Prize [17] as a quality paper. Taguchi method extended outward slowly, and many manufacturers in the world also expected to adopt the Taguchi method to improve the defect rate, such as car manufacturer (Ford), mobile phone manufacturer (Motorola), camera manufacturer (Kodak) and so on.

The adjustment for multi-level factors includes the following four methods:

(1) Trial-and-Error

If it is not easy to get the right answer directly when solving a problem, the method of probability can be used to try to change the problem that needs a solution.

(2) One-Factor-at-a-Time Experiment

Only one factor will change every time during the experiment, while other factors remain the same. It is possible to explore the effects of change of each factor.

(3) Full-Factorial Experiment

In the same experiment, all factors are permutated and combined so that each factor will combine with other factors in the same experiment. If there are more factors, it will be time-consuming.

(4) Orthogonal Array

The orthogonal array adopts a systematic way to permute and combine to ensure the interaction among factors. However, unlike the full factorial experiment, there is no need to actually implement each experiment. The orthogonal array has improved the reliability of the experiment, reaching accuracy within a short period of time.

III. RESEARCH METHODOLOGIES

A. Research Process

Fig. 1 illustrated the entire experimental process and framework, with a five-step analysis of how this study aimed to construct a real estate appraisal model.

Step 1. Data sources were obtained from batch data of actual

price registration for real estate transaction, Ministry of Interior. Actual price registration data from every township in Taichung City in February 2014 were collected.

Step 2. A lot of data were not complete. Data were examined individually to delete those that were missing a variable. As for the development of special cases, data of malls or public construction openly built by the government such as parks were seen as outliers and were deleted since they were not transactions done by ordinary consumers.

Step 3. When collecting data, raw data were maintained if possible, so there were many dimensions. But some were not helpful in terms of building a predictive model. Therefore, the Information Gain and Gain Ratio were used to find the dimensions that had more significant impact on building a model. For one thing, accuracy can be enhanced, and for another, efficiency of modeling can be improved and time cost can be reduced.

Step 4. Different machine learning was used to build a predictive model. This study used the artificial neural network with Taguchi method for parameter optimization and compared the differences among the comparison tree, the support vector machine and the linear regression.

Step 5. Different evaluation indicators were used to analyze the predictive ability after models were built by different machine learning.

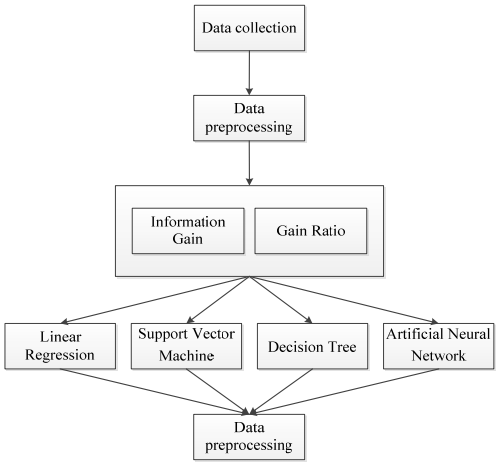

Fig. 1 Working concepts and procedures

B. Description of Dataset

Data in this research were obtained from the real estate actual price registration of the Ministry of Interior in February 2014. After data preprocessing, 2690 data remained [18].

C. Selection of Dimensions

Dimension selection is also called variable selection, using information gain and gain ratio in information theory to select dimensions. The two methods are introduced in the following.

TABLE III
DESCRIPTION OF DATA SET

Item	Description
Township	Data were collected from every district of Taichung City.
The total area of land with transfer of ownership	The area of land with transfer of ownership at the time of the transaction of buildings or lands
Land use zoning	Land use is divided into residential land, commercial land or agricultural land
Transfer level	Single floor, the entire building, the underground, arcade and floors
Total number of floors	Up to 39 floors of high-rise buildings.
Building Type	1. Apartment (under five floors(inclusive) with no elevator) 2. Residential buildings (more than 11 floors(inclusive) with elevator) 3. Stores (Shops) 4. Suites (with one bedroom, one dining room, and one bathroom) 5. Townhouses 6. Condominiums (under 10 floors(inclusive) with elevator) 7. Office and commercial buildings
Main purpose	The main purpose of the building is divided into residential, commercial or residential and commercial combined.
Main building materials	1. See other registered items 2. Steel structure 3. Steel reinforced concrete 4. Reinforced concrete
Total area of building transfer	Total usable areas in buildings
Whether there is management organization	Whether the building has a good management system
Parking categories	1. First floor flat 2. Lifting plane 3. Lifting machinery 4. Ramp plane 5. Ramp machinery 6. Other
Price per square meter	Selling price per square meter

(1) Information Gain

Information gain variable selection method was published by Quinlan, a famous decision tree scholar, in 1979[19] .Mainly, the method assesses the extent of the amount of information of each variable by calculating the size of the amount of information. The calculation index of the amount of information is called entropy. Entropy is used to calculate impurity of overall dataset in computer science. When the dataset has higher impurity, the value of entropy is close to 1. The following formula explains how to calculate entropy.

$$Entropy(S) = -p_a \log_2 p_a - p_b \log_2 p_b \quad (1)$$

$$\text{Entropy}(S) = \text{Data set } S$$

$$p_a = \text{variable } a$$

$$p_b = \text{variable } b$$

Information gain can be used to assess the importance of different variables. The greater the information gain value, the more influential this variable will be on classification of dataset. When the information gain value is at its maximum, it will be used as root. The following equation describes the calculation of information gain

$$Information\ Gain(S, A) = Entoypy(S) - \sum_{y=1}^{x} \frac{|S_y|}{S} Entyopy(S_y) \quad (2)$$

$$Information\ Gain(S,A) = Data\ set\ S\ for\ variableA$$
$$\sum_{y=1}^{x} \frac{|S_y|}{S} Entyopy(S_y) = VariableAfor\ data\ x$$

(2) Gain Ratio

Gain ratio variable selection method is also presented by scholar Quinlan in 1986 [20]. It was developed from information gain. The information from information gain is normalized, with the main purpose to reduce deviation correction problems of information gain during dimension selection.

$$Gain\ Ratio(S,A) = \frac{Information\ Gain(S,A)}{Entropy(S,A)} \quad (3)$$

D.Artificial Neural Network Architecture

Input layer: Refers to the most suitable real estate transaction information that is picked after dimension selection and used as input variable.

Hidden layer: The main function is to receive input data. Through self-learning and adjustment, it can deal with non-linear problems. The hidden layer is usually one or two layers. Too many layers may cause over-learning. As a result, it depends on complexity of the problem.

Output layer: The output layer is the final result of the predictive output. The final output of the study is the predictive real estate transaction price.

Neuron: The number of neurons in the study is usually a combination of the input layer and the output layer [21].

1. (Input layer processing unit number +output layer processing unit number)
2. (Input layer processing unit number +output layer processing unit number)/2
3. (Input layer processing unit number +output layer processing unit number) *2

Learning Rate: The method of weight is used to increase the ability of artificial neural network in handling nonlinear problems. If the value is too small, it is not conducive to learning. If the value is too large, there will be excessive learning problems. Hence, this study used a systematic way to allocate, typically between 0.1 and 1.0.

Momentum: The method of weight is used to move beyond original local training. If the momentum is too small, it will not move beyond to forma local optimal solution. Yet if the momentum is too big, the training results will be affected and the optimal solution won't be found easily.

Epochs: It is usually the number of times of the experiment. There will be results only when the experiment is executed to a certain extent. If the number of epochs is too small, it is possible that the experiment stops before reaching convergence. If the number is too big, it will lead to over-learning easily, resulting in divergence after reaching convergence. Therefore, an appropriate number of epoch is needed.

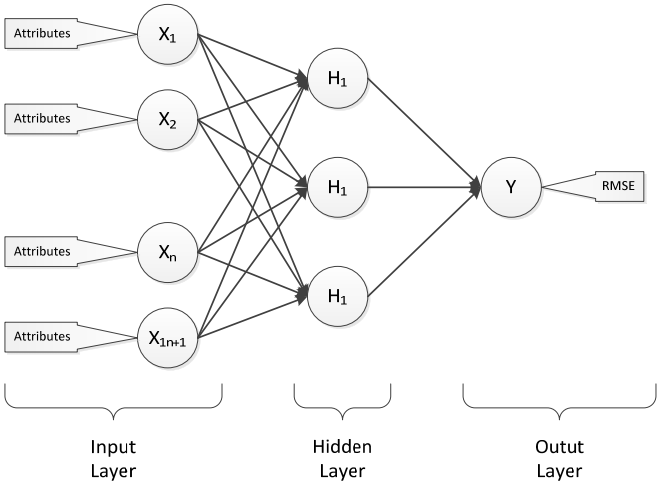

Fig. 2 Artificial Neural Network topology

E. Taguchi Method

The main purpose of using Taguchi method in this study is to optimize the parameter setting of artificial neural network. If the main four parameters - neurons, momentum, learning rate and epochs - are not planned properly, it is possible that a predictive model for real estate transaction prices will not be built. This study used Minitab as the software to program Taguchi method. The selection criteria of Taguchi method used in this study are described in the following:

1. Selection and Appropriate Level of Experimental Factors

This study used artificial neural network to establish valuation for real estate transaction. Scholars at home and abroad focused on the four major experimental factors in the artificial neural network, namely, neurons, learning rate, momentum, and epochs, so Taguchi method selects this sub-variable as an experimental factor; different levels are primarily used as a range for planning variables. If the gap between variable segments is relatively large multi-levels can be set.

2. Selection of Orthogonal Array

The main purpose of orthogonal array is to plan how many experiments are needed. If the result of each experiment is considerably different, more experiments can be set up. If the experimental results do not change much, the number of experiments can be reduced to save time. Most researchers will adopt rules of thumb to select the appropriate orthogonal array to conduct Taguchi analysis.

3. S/N Ratio (Signal to Noise Ratio)

The main objective of S/N ratio is to utilize the unique analysis of Taguchi method to conduct valuation. Each experiment planned by the orthogonal array will be analyzed on the S/N ratio to assess the impact of different level in each experiment. The S/N ratio can be divided into three types: the larger the better (LTB), the smaller the better (STB) and nominal the best (NTB). The purpose of S/N ratio is decided depending on the experiment needs.

F. Support Vector Machine

In 1995, American scholar Vapnik developed a method called support vector machines (SVM) to solve complex classification problems [22]. SVM is of supervised learning and is able to find the best classification method and build a model after learning. The way to build SVM is to separate data that is classified and then project them onto a multi-dimensional space, called hyperplane. After a lot of data training, data are separated into different categories to establish a model, which can be used to predict unsorted data.

Core function in LibSVM is used to project data onto the hyperplane. When the data exceeds two dimensions, the core function will project onto a high-dimensional space. When using the concept of hyperplane to distinguish, the core function is required to converse different data types. Therefore, different core function can be used to split data. The core function was first proposed by Aizerman, Braverman, and Rozoner in 1964 for selecting characteristics. Selecting the appropriate and effective core function is very important, which is the only indispensable step. Common core function is described below:$\mathrm{K}(u_i, v_j)$

1. Linear Function:

$$\mathrm{K}(u_i, v_j) = u_i * v_j \tag{4}$$

2. Polynomial Function:

$$\mathrm{K}(u_i, v_j) = (gamma * u_i * v_j + coef0)\text{^}degree \tag{5}$$

3. Radial Basis Function:

$$\mathrm{K}(u_i, v_j) = exp(-gamma * |u_i - v_j|\text{^}2) \tag{6}$$

4. Sigmoid Function:

$$\mathrm{K}(u_i, v_j) = tanh(gamma * u_i * v_j + conf0) \tag{7}$$

Scholars at home and abroad have done further research on the core function of SVM. Many scholars also explored radial basis function (RBF), also called Gaussian function. This study also conducted experiments using a Gaussian function. The following are the disadvantages and disadvantages as well as application of Gaussian function found by [23]-[24]:

1. Gaussian function plays an important role in many algorithms, and is useful in nonlinear data processing.
2. It is easy to operate, thus reducing operational difficulty and having good predictive ability.
3. When the variable is entered, data will be normalized to between 0 and 1.After narrowing down the range of data, the distortion problem caused by large differences among variables can be effectively reduced.

G.Linear Regression

Linear regression can be regarded as a statistical method to identify causal relations from the database. Pindyck and Rubinfeld mentioned that there is a dependent variable and one or more independent variables. Minimum sum of square of linear regression equation is used to find out the rules and to establish an association between the dependent variable and the independent variable, called linear regression model [25] . The reason why linear regression is powerful is because it can use data in the database to conduct causal analysis to further reach the predictive effect. In terms of prediction, linear regression can integrate multiple variables for analysis, and the factor can operate independently without being affected by other factors, which will lead to distorted predictive results made by the model.

H.Decision Tree

Decision Tree has a long history of development, and can also be called classification tree. When it is necessary to use classification rules, decision tree will definitely come into people's mind. The principle of decision tree is easy to understand and it does not require too much statistical reasoning. One may conduct classification according to the objective of analysis, and find relevant characteristics from classified dataset. By using the eigenvalue, a systematic classification rule will be developed after calculation. The "if ... than" rule is then developed. When new samples or test datasets need to be classified, this rule tree can be used. According to the eigenvalues in the rule tree that has been established, one may predict how unsorted data can be classified into appropriate category. In the past, many scholars have studied and improved decision tree algorithm, which consists of the root node, branch and the leaf node. When establishing a model, it is necessary to split a certain percentage of the training datasets as the establishment basis. Three classic representatives of decision tree are introduced in the following: ID3 (Iterative Dichotomiser 3), CART (Classification and Regression Trees), and C4.5 algorithm [26]-[28]:

I. Evaluation Index

After a series of experiments obtain results, it is necessary to rival with different methods. An objective assessment criterion is needed. In the past, scholars often used the following four indicators to assess the error rate: root mean square error (RMSE), mean absolute error (MAE), relative absolute error (RAE), and root relative squared error (RRSE).

(1) Root Mean Squared Error, RMSE

The root mean squared error focuses on the same basis when doing comparison to find the difference between predictive value and the actual value. Therefore, it is better that the root mean squared error is small.

$$\mathrm{E} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)^2} \tag{8}$$

$$P_i = Predicted\ i$$

$$A_i = Actual\ i$$

$$\mathrm{n} = \text{Total number of data}$$

(2) Mean Absolute Error, MAE

The mean absolute error mainly aims at observing the difference between the predictive value and the actual value. In deviation measurement, the mean absolute error is often used as the evaluation criteria. Therefore, it is better than the mean absolute error is small.

$$E = \sum_{i=1}^{n} \left| \frac{A_i - P_i}{A_i} \right| \quad (9)$$

$P_i = Predicted\ i$

$A_i = Actual\ i$

n = Total number of data

(3) Relative Absolute Error, RAE

The relative absolute error mainly aims at assessing the difference between the predictive value and the actual value. Before calculation, the error will be normalized. For example, different errors will be normalized between 0 and 1 to facilitate comparison.

$$E_j = \frac{\sum_{i=1}^{n} |P_{ij} - A_i|}{\sum_{i=1}^{n} |A_i - A_m|} \quad (10)$$

$P_{ij} = Data\ set\ j\ for\ predicted\ i$

$A_i = Actual\ i$

n = Total number of data

$A_m = Average\ for\ A_i$

(4) Root Relative Squared Error, RRSE

The main aim of the root relative squared error is to assess the average error between the predictive value and the observed value. Before calculating the average error, it is necessary to have the same variable as a baseline.

$$E_j = \sqrt{\frac{\sum_{i=1}^{n} (P_{ij} - A_i)^2}{\sum_{i=1}^{n} (A_i - A_m)^2}} \quad (11)$$

$A_i = Actual\ i$

n = Total number of data

$A_m = Average\ for\ A_i$

IV. Experiential Design and Results

A. Research Process

This study focused on current leading economic indicator, that is, real estate transaction prices. The sources of information were from the ROC Ministry of the Interior Information Network. The data were about Taichung City. Since Taichung City and Taichung County merged together to upgrade to a municipality, many urban renewal plans are taking place and many business owners or investors have already stepped into local real estate market. In recent years, many reports revealed cases of "fake registration, real market pushing" and "Covering up for low prices by registering high prices." Such unlawful operation has seriously impacted the general public to purchase their homes. The government and business have conducted investigation and other methods to check whether the trading market prices were deliberately manipulated. This research selected real estate transaction actual price registration batch data in February 2014 and used artificial neural network to build a real estate transaction predictive model. Taguchi method was also used to effectively regulate artificial neural network parameter setting. The study used support vector machines (SVM), linear regression (LR) and decision trees (DT) to verify if the proposed method can effectively improve the quality of modeling.

B. Dimension Selection Results

1. Information Gain

Information gain in WEKA variable selection was used. The results in Table IV show execution results of information gain. After all variables were ranked, predictive analysis was performed. RMSE was used to decide whether to delete the variables or not. The bottom ones on the ranking list would be deleted first. After deleting one variable, the remaining variables were re-tested and RMSE was recorded. When the experiment was about the "urban land zoning," the RMSE reached its lowest because only variables containing 9 were retained.

TABLE IV
Information Gain Variable Selection

Attribute	Rank	Information Gain	RMSE
Floor-total	1	0.50749	0.3027
Township	2	0.46051	0.246
Floor	3	0.23617	0.2472
Land_area	4	0.15904	0.2428
Building-area	5	0.1478	0.2424
Main_purpose	6	0.11641	0.2437
Parking-type	7	0.09227	0.2411
Building-type	8	0.05665	0.2416
Land_Classification	9	0.05041	0.2371
Build-material	10	0.04111	0.238
Manage	11	0.00389	0.239

2. Gain Ratio

Gain ratio in WEKA variable selection was used. The results in Table V show execution results of gain ratio. After all variables were ranked, predictive analysis was performed. RMSE was used to decide whether to delete the variables or not. The bottom ones on the ranking list would be deleted first. After deleting one variable, re-tested the remaining variables and recorded RMSE. When the experiment was on the "parking categories," the RMSE reached its lowest because only variables containing 8 were retained.

TABLE V
Gain Ratio Variable Selection

Attribute	Rank	Gain Ratio	RMSE
Build-material	1	0.2389	0.357
Floor-total	2	0.1522	0.304
Township	3	0.1484	0.251
Building-area	4	0.1237	0.246
Land_area	5	0.1198	0.240
Main_purpose	6	0.1186	0.242
Building-type	7	0.0778	0.240
Parking-type	8	0.0697	0.236
Land_Classification	9	0.062	0.237
Floor	10	0.056	0.238
Manage	11	0.0246	0.239

C. Artificial Neural Network with Taguchi Method

1. Experimental Factors and Standard Setting

Table VI illustrates setting for selecting experimental factors and levels. The following four factors were set as the main parameter setting factors for the artificial neural network:

neurons, learning rate, momentum, and epochs. These four factors have a significant impact on model establishment. If parameters needed for the dataset are not optimized, the model that is built will not predict accurately; To select three levels for factor design, only four main factors need to be adjusted. Therefore, there is no need to consider the interaction among factors when choosing the three-level standard.

TABLE VI
EXPERIMENTAL FACTORS AND LEVELS

Factor / Level	Neuro	Lr	Mt	Epochs
Level 1	3	0.1	0.1	3000
Level 2	6	0.5	0.5	6000
Level 3	12	0.9	0.9	10000

D. Deciding Experimental Methods and Factor Configuration

This study selected $L_9(3^4)$ orthogonal array as experimental planning. Table VII illustrates the combination of factors in each experiment. There were a total of three levels, four factors and nine experiments. Orthogonal array planning can have fewer numbers of experiments but fully adjusting results. The following experiment is based on the root mean square error (RMSE), so STB signals and S/N ratio formula were chosen for analysis. The purpose is to gain smaller error in every experiment and to get a higher weight.

TABLE VII
L_9 ORTHOGONAL ARRAY EXPERIMENT RESULTS

No.	Neuros	Lr	Mt	Epochs	S/N Ratios
1	3	0.1	0.1	3000	20.6839
2	3	0.5	0.5	6000	19.4730
3	3	0.9	0.9	10000	20.5043
4	6	0.1	0.5	10000	20.2253
5	6	0.5	0.9	3000	20.1651
6	6	0.9	0.1	6000	20.6007
7	12	0.1	0.9	6000	20.5147
8	12	0.5	0.1	10000	20.8753
9	12	0.9	0.5	3000	20.0661

Fig. 3 shows the main effect plot of main parameters of the artificial neural network, which is based on S/N ratio. The smaller the S/N ratio, the lower the root mean square error (the higher the model prediction accuracy). Fig. 3 demonstrates the optimized setting conditions for the artificial neural network parameters after optimization. When creating the model, the number of neurons is set at 3, the learning rate 0.5, momentum 0.5, and the number of epochs 6000. Thus, a predictive model for real estate transaction was built.

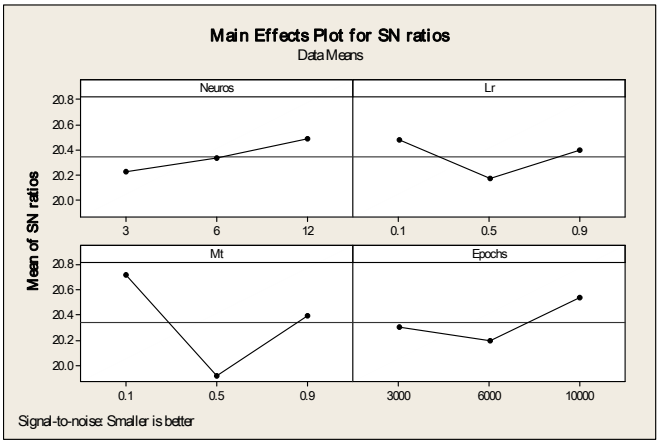

Fig. 3 The optimal parameters among factors

Table VIII illustrates different S/N ratio among factors. The difference is established by comparing the highest and the lowest ration of the same factor. The biggest difference is ranked as the first one, and the smallest difference is ranked as the fourth one. Number one in the ranking represents the most significant effect in building a predictive model.

TABLE VIII
DIFFERENCES AMONG DIFFERENT LEVELS

Level	Neuros	Lr	Mt	Epochs
1	20.22	20.47	20.72	20.31
2	20.33	20.17	19.92	20.2
3	20.49	20.39	20.39	20.53
Difference	0.26	0.3	0.8	0.34
Rank	4	3	1	2

E. Decision Tree

This study used different algorithms in WEKA software. Three decision tree algorithms--ID3, J48, CART--were selected. They can be applied to different data types, so they are the most commonly used three algorithms by researchers nowadays. Tables IX and X present the experimental results.

TABLE IX
DIFFERENT DECISION TREES WERE EXPERIMENTED USING 80:20 DATA DISTRIBUTION

Evaluation index	ID3	J48/C4.5	CART
Data distribution	80:20	80:20	80:20
RMSE	0.2239	0.2223	0.2242
MAE	0.089	0.0927	0.0928
RAE(%)	63.43%	64.45%	64.46%
RRSE(%)	84.79%	83.01%	83.73%

TABLE X
DIFFERENT DECISION TREES WERE EXPERIMENTED USING 10-FOLD DATA DISTRIBUTION

Evaluation index	ID3	J48/C4.5	CART
Data distribution	10-fold	10-fold	10-fold
RMSE	0.2187	0.2203	0.2197
MAE	0.0879	0.0927	0.0922
RAE(%)	62.29%	64.32%	64.02%
RRSE(%)	82.48%	82.12%	81.90%

F. Comparison of Different Methods

Tables XI and XII conduct the different experiments and show the experimental results by using artificial neural network (ANN), support vector machines (SVM), linear regression (LR) and decision trees (ID3, J48, CART) approaches. Two different proportions of training datasets and test datasets were selected to improve the reliability. The evaluation index is the root mean square error (RMSE). The following experimental results were presented.

TABLE XI
DIFFERENT MACHINE LEARNING RESULTS WERE EXPERIMENTED USING 80:20 DATA DISTRIBUTION

80:20 (RMSE)	ID3	J48	CART	ANN	SVM	LR
Info Gain	0.2225	0.2036	0.2135	0.164	0.2369	0.1298
Gain Ratio	0.2128	0.2056	0.2147	0.2106	0.2416	0.1369

TABLE XII
DIFFERENT MACHINE LEARNING RESULTS WERE EXPERIMENTED USING 10-FOLD DATA DISTRIBUTION

10-fold (RMSE)	ID3	J48	CART	ANN	SVM	LR
Info Gain	0.2172	0.2012	0.2081	0.1249	0.2371	0.1396
Gain Ratio	0.2091	0.2082	0.2095	0.314	0.236	0.145

1. The decision trees include ID3, J48 and CART. There is small difference between dimension selection of information gain and gain ratio. Therefore, these two selection methods have selected the best variable for the decision tree.
2. In terms of artificial neural network (ANN), information gain was used as variable selection to achieve the best predictive ability of this study. The RMSE was the lowest when using 10-fold experiment, which is more suitable than gain ratio.
3. In terms of support vector machine (SVM), there is little difference between using information gain and gain ratio as dimension selection. Therefore, using information gain and gain ratio will reach the same effect.
4. In terms of linear regression (LR), when using information gain and gain ratio as dimension selection, information gain have higher accuracy. Therefore, it is recommended to use information gain dimension selection for linear regression.

V. CONCLUSION

This paper studied parameter setting for artificial neural network. More efficient way was used to adjust the four main parameters in artificial neural network. Therefore, Taguchi method, a common method used in industrial quality engineering, was applied. A suitable orthogonal array was used as base to adjust the parameters. Based on the advantages of artificial neural network, highly complex issues can be handled. When the artificial neural network receives different variable inputs, it can recommend the best parameter setting, and can self-learn and adjust the best causal relational model. As a livelihood issue, real estate transaction has been discussed by scholars at home and abroad. Real estate is highly sensitive, and many factors affect the price of real estate transaction, so it is called a highly complex issue. The study adopted several commonly used machine learning algorithms to conduct experiments and finally from the experimental results, it was concluded that the use of artificial neural network with the Taguchi method had a more significant predictive ability. The conclusions were summarized as the following three points:

1. This study used back-propagation neural network (BPNN) to build a real estate transaction predictive model. BPNN is a kind of supervised learning. After the input layer received the dataset, it would allocate them to the hidden layer to calculate learning rate and value-added weight. After repeated calculation and learning, the outcome was reached and a predictive model for the real estate transaction price was established.
2. Taguchi method was used to effectively adjust the four main parameters in back-propagation neural network to calculate permutations and combinations of different levels in L_9 orthogonal array. The nine experiments selected the best parameter combination based on S/N ratio calculation and recommendation, namely, optimized parameter for back-propagation neural network.
3. Different dimension selection tools were used to pick out different variables. Related variables from the dataset and real estate transaction prices were sorted and unnecessary variables were deleted to avoid affecting the predictive effects in subsequent experiments. From the perspectives of different machine learning methods, information gain and gain ratio had their respective advantages and disadvantages. Therefore, it is very important to use correct dimension selection for different machine learning.

ACKNOWLEDGEMENT

The authors thank the support of Ministry of Science and Technology (MOST) of the Republic of China (ROC) to this work under grant no. MOST103-2410-H-025-022-MY2.

REFERENCES

[1] J.F.C. Khaw, B.S. Lim, and L.E.N. Lim, "Optimal design of neural networks using the Taguchi method," *Neurocomputing*, vol.7, no.3, pp. 225-245, 1995.
[2] A. Tortum, N. Yayla, C. Çelik, and M. Gökdağ, "The investigation of model selection criteria in artificial neural networks by the Taguchi method,"*Physica A: Statistical Mechanics and its Applications*, vol.386, no.1,pp. 446-468, 2007.
[3] W.C. Chen, Y.Y. Hsu, L.F. Hsieh, and P.H. Tai, "A systematic optimization approach for assembly sequence planning using Taguchi method, DOE, and BPNN," *Expert Systems with Applications*, vol.37, no.1, pp. 716-726, 2010.
[4] K.Y. Chang, "The optimal design for PEMFC modeling based on Taguchi method and genetic algorithm neural networks," *International Journal of Hydrogen Energy*, vol.36, no.21, pp. 13683-13694, 2011.
[5] J.R. Jung, and B.J. Yum, "Artificial neural network based approach for dynamic parameter design," *Expert Systems with Applications*, vol.38, no.1,pp. 504-510, 2011.
[6] T.B. Asafa, N. Tabet, and S.A.M. Said, "Taguchi method–ANN integration for predictive model of intrinsic stress in hydrogenated amorphous silicon film deposited by plasma enhanced chemical vapour deposition," *Neurocomputing*, vol. 106, pp. 86-94, 2013.
[7] C.O. Chang. Real estate in the world: trade, investment, agency, policy, Yuan-Liou Publishing, 1990.

[8] W.J. McCluskey, and A.S. Adair, *Computer Assisted Mass Appraisal*. 1997.
[9] F. Rosenblatt, "The perceptron - a perceiving and recognizing automaton," Cornell Aeronautical Laboratory report, pp.85-460-1, 1957.
[10] W.S. McCulloch, and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol.5, no.4, pp. 115-133, 1943.
[11] D.O. Hebb, *The organization of behavior: a neuropsychological theory*. John Wiley and Sons, 1949.
[12] M. Minsky, and S.Papert, *Perceptrons*. MIT Press. Cambridge, Ma, 1969.
[13] S. Grossberg, "Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors," *Biological Cybernetics*, vol.23, no.3,pp. 121-134, 1976.
[14] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, vol.43, no.1,pp. 59-69, 1982.
[15] J.J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," Proceedings of the National Academy of Sciences, vol.81, no.10, pp. 3088-3092, 1984.
[16] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning representations by back-propagating errors," in Neurocomputing: foundations of research, A.A. James and R. Edward, Editors. MIT Press. pp. 696-699, 1986.
[17] H.H. Lee, *Taguchi Methods: Principles and Practices of Quality Design*. GauLih Book Publishing (4nd), 2013.
[18] Ministry of the Interior of R.O.C., Monthly Bulletin of Interior Statistics, 2014. Available at: http://sowf.moi.gov.tw/stat/month/list.htm.
[19] J.R. Quinlan, "Discovering rules by induction from large collections of examples," *Expert Systems in the Microelectronic Age*, pp. 168-201, 1979.
[20] J.R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no.1, pp. 81-106, 1986.
[21] I.C. Yeh, *Neural Network Model: Application and Implementation* (9rd), Scholars Book Publishing, 2009.
[22] C. Cortes, and V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol.20, no.3, pp. 273-297, 1995.
[23] S.S. Keerthi, and C.-J. Lin, "Asymptotic Behaviors of Support Vector Machines with Gaussian Kernel," *Neural Computation*, vol.15, no.7, pp. 1667-1689, 2003.
[24] C.W. Hsu, C.C. Chang, and C.J. Lin, A practical guide to support vector classification. 2003.
[25] R.S. Pindyck, and D.L. Rubinfeld, *Econometric models and economic forecasts*. Boston, Mass.: Irwin/McGraw-Hill, 1998.
[26] J. Mingers, "An empirical comparison of selection measures for decision-tree induction," *Machine Learning*, 1vol.3, no.4, pp. 319-342, 1989.
[27] U.M. Fayyad, and K.B. Irani, "On the handling of continuous-valued attributes in decision tree generation," *Machine Learning*, vol.8, no.1, pp. 87-102, 1992.
[28] T.G. Dietterich, "Ensemble Methods in Machine Learning, in Multiple Classifier Systems," *Lecture Notes in Computer Science*, vol. 1857, pp. 1-15, 2000.

Dr. Chen is an Associate Professor of Information Management at National Taichung University of Science and Technology, Taiwan. His current research interests include artificial intelligent, soft computing, bio-inspired computing, financial engineering, and data mining. Dr. Chen's research is published or is forthcoming in Information Sciences, Applied Soft Computing, Journal of Educational Technology & Society, Journal of Information Science, The Electronic Library, Computers and Mathematics with Applications, Quantitative Finance, Expert Systems with Applications, Soft Computing, and a number of national and international conference proceedings.

Dr. Fan is an Associate Professor of Information Management at National Taichung University of Science and Technology, Taiwan. His current research interests include artificial intelligent and financial engineering. Dr. Fan's research is published or is forthcoming in Neural Computing and Applications, Journal of Information Science and Engineering, IEICE Transactions, and a number of national and international conference proceedings.

Dr. Chen is an Assistant Professor of Management Information Systems at National Chung Hsing University, Taiwan. Her current research interests include RFID, context awareness, wireless and sensor network, e-learning, and smart living. Dr. Chen's research is published or is forthcoming in Information Sciences, Information Fusion, Computers and Education, Journal of Educational Technology & Society, The Electronic Library, International Journal of Mobile Communications, Personal and Ubiquitous Computing, Expert Systems with Applications, International Journal of Information Technology and Management, and a number of national and international conference proceedings.

Mr. Jhong is a graduate student of Information Management at National Taichung University of Science and Technology, Taiwan. His current research interests include artificial intelligent and financial engineering.