[ED03] Automated low cost house demand forecasting for urban area

Noor Yasmin Zainun[1], Muhd Zaimi Abd. Majid[2]

[1]Faculty of Environment and Structural Engineering, Department of Building and Construction, Kolej Universiti Teknologi Tun Hussein On, 86400 Parit Raja, Batu Pahat, Johor, Malaysia.
[2]Faculty of Civil Engineering, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia.

Introduction

Historically, urbanization was the product of industrial expansion and rapid economic growth. In developing country however, the process has been characterized by rapid urban growth without corresponding economic growth of the cities. This resulted in the emergemve of the informal sector and squatter settlements. In Malaysia, urbanization is expexted to continue, with the Department of Statistics projecting the urban population at 64% of total population by the year 2020. Due to the increment of the demand for houses especially in urban area is very significant and vital, there is increasing need to produce an accurate demand forecast for housing. Recently, neural networks have been successfully applied to many applications in resolving civil engineering (Cheu *et. al.* 1991; Moselhi *et. al*. 1991) and structural engineering problems (Furuta *et. al*. 1991; Hajela *et. al.* 1991) as well as in many other fields (Weigend *et. al*. 1990; Karunanithi *et. al.* 1992). A neural network is a computational method inspired by studies of the brain and nervous systems in biological organisms. Typically, a neural network consists of a set of layered processing units and weighted interconnections. One of the well-known classes of neural networks used for forecasting application is a feed-forward network. In a feed-forward network the weighted connections feed activities only in the forward direction from the input layer to the output layer. The most commonly used training algorithm for feed-forward networks is the back-propagation algorithm. The back-propagation algorithm is a gradient descent method in which weights of the connections are updated using partial derivatives of error with respect to weights.

The objective of this paper is to forecast demand on low cost housing in urban area in Malaysia using Artificial Neural Network (ANN) approach. This study focused on 2 districts of Selangor, Malaysia, including Petaling and Gombak, among the areas that have reached the highest urbanized level in the country. Then the computerized model is developed so that unskilled user also can used to forecast with only few minutes.

Methodology

The methodology of this study including series of trial and error process to find out the optimum number of training and testing data, finding out the suitable number of hidden neurons, learning rate, and momentum rate for the network and screening the result using the best Neural Network (NN) model in forecasting demand on low cost housing in Petaling and Gombak. The computerized modules internally interacted with each other by multiple Document Interface (MDI) Technique. SPSS and NeuroShell2 packages are internally referenced by Object Linking and Embedding (OLE).

Independent and dependent variables

To perform NN model, the important or significant of the independent variables should be determined to avoid longer generalization (Khairulzan, 2002). In this study, Principal Components Analysis (PCA) is used to derive new variables; that is the significant variables from the nine selected indicators using SPSS 10.0 package. The indicators are: (1) population growth; (2) birth rate; (3) average mortality baby; (4) unemployment rate; (5) inflation rate; (6) gross domestic product (GDP); (7) poverty rate; (8) income rate; and (9) housing stock. The dependent variable is the monthly time series data on demand on low cost housing each district for 5 years from February 1996 to November 2000. From the analysis, it was find out that Petaling having 2 Principal Components (PCs) while Gombak having 3 PCs. These PCs become the inputs in the NN model and time series data on housing demand become the actual output.

Training and testing data

All data are divided into two categories for training and testing. Since training and testing data will randomly selected, a series of trial and error process has been carried out to find the optimum number of training and testing data. Training and testing performance are evaluated using linear correlation coefficient, r.

$$r = \frac{SS_{XY}}{\sqrt{SS_{XX} \cdot SS_{YY}}}$$

Where

$$SS_{XY} = \sum XY - \frac{(\sum X)(\sum Y)}{n}$$

$$SS_{XX} = \sum X^2 - \frac{(\sum X)^2}{n}$$

$$SS_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$$

Where;n = the number of patterns,
X = set of actual outputs, and
Y = predicted outputs.

From the process, the optimum training and testing data for Petaling are 40 and 9 with highest r = 0.6874 while for Gombak are 46 and 3 with highest r = 0.8105 respectively as shown in Table 1.

Creating Neural Network model

According to Cattan (1994), a network is required to perform two task; (1) reproduce the patterns it was trained on and (2) predict the output given patterns it has not seen before, which involves interpolation and extrapolation. In order to perform these tasks, a backpropagation network with one hidden neutron is used. Sigmoid functions are used as the activation function for hidden and output layers. To find out the best number of hidden neurons for the network, the default setting of backpropagation algorithm in Neuroshell 2 is applied, where the value for learning rate and momentum is 0.1. The average error used is 0.001 and 20,000 learning epochs. The numbers of input nodes for Petaling are two since they having two principal components (PCs) as the input variables while the number of input nodes for Gombak are three since Gombak having three PCs as the input variables. The number of output neuron for the three districts is one since the output is only one that is the housing demand. Figure 1 shows the neural network topology with one hidden layer for Petaling.

TABLE 1 Optimum training and testing data for Petaling and Gombak

District	Training	Testing	r
Petaling	40	9	0.6874
Gombak	46	3	0.8105

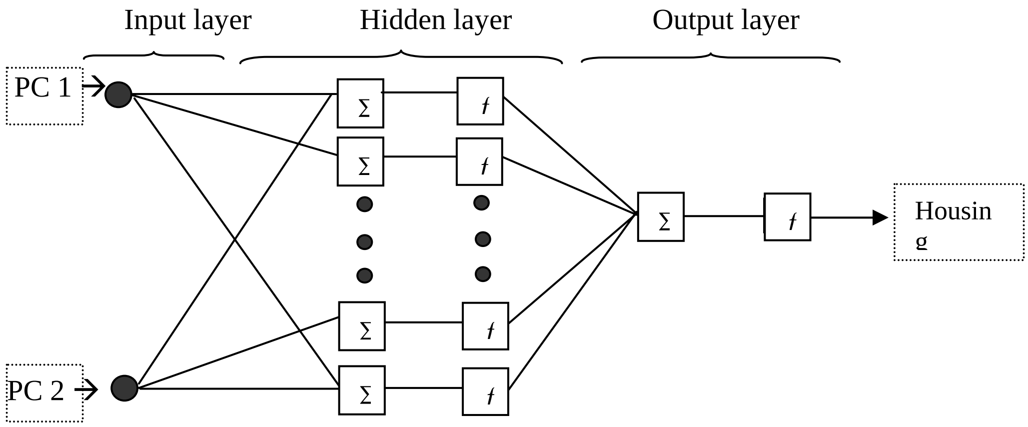

FIGURE 1 Neural Network topology with 2 inputs for Petaling

From Figure 2, the training graph shows that the networks performed well where the value of linear correlation coefficient, r are above 0.5. The networks performed uniformly between r = 0.5913 to r = 0.6091. For testing, the networks performance are up and down between r = 0.4635 to r = 0.6874. The highest linear correlation coefficient is obtained using 8 numbers of neurons in hidden layer. It shows that the best Neural Network to forecast demand on low cost housing in Petaling is 2-8-1, which is 2 numbers of neurons in input layer, 8 numbers of neurons in hidden layer and 1 number of neuron in output layer.

Figure 3 shows that most of the networks performance of training are very good where value of r are uniformly between 0.6092 to 0.7403 except using 2, 3, 4, 13, 16, and 28 numbers of neurons where the value of r is only between 0.1612 to 0.2996. The lowest network performance is using 2 numbers of neurons and the highest network performance is when using 35 numbers of neurons. The testing graph also shows that most of networks performed very good where value of r is very high that is between 0.7671 to 0.8105 except using 2, 3, 12, 13, 14, 16, 22, 28, 30, and 34 numbers of neurons. The lowest network performance is when using 14 numbers of neurons while the highest network performance is when using 11 numbers of neurons. Therefore, the best Neural Network to forecast demand on low cost housing in Gombak is using 3-11-1, which is 3 numbers of neurons in input layer, 11 numbers of neurons in hidden layer and 1 number of neuron in output layer.

After that, another series of trial and error processes are done by changing the learning rate and momentum rate. It was found out the best NN to forecast demand on low cost housing for Petaling and Gombak are using 0.7 learning rate and 0.4 momentum rate with r value of 0.6796 and 0.8162 respectively. Lastly, the demand on low cost housing for Petaling and Gombak is forecasted for September 1999 and October 1999 using their best Neural Network. Table 2 and 3 below show the actual and forecasted demand by Neural Network for Petaling and Gombak. Mean absolute percentage error, MAPE is calculated to evaluate the forecasting performance. According to Sobri Harun (1999), the ability of forecasting is very good if MAPE value is less than 10% while MAPE for less than 20% is good.

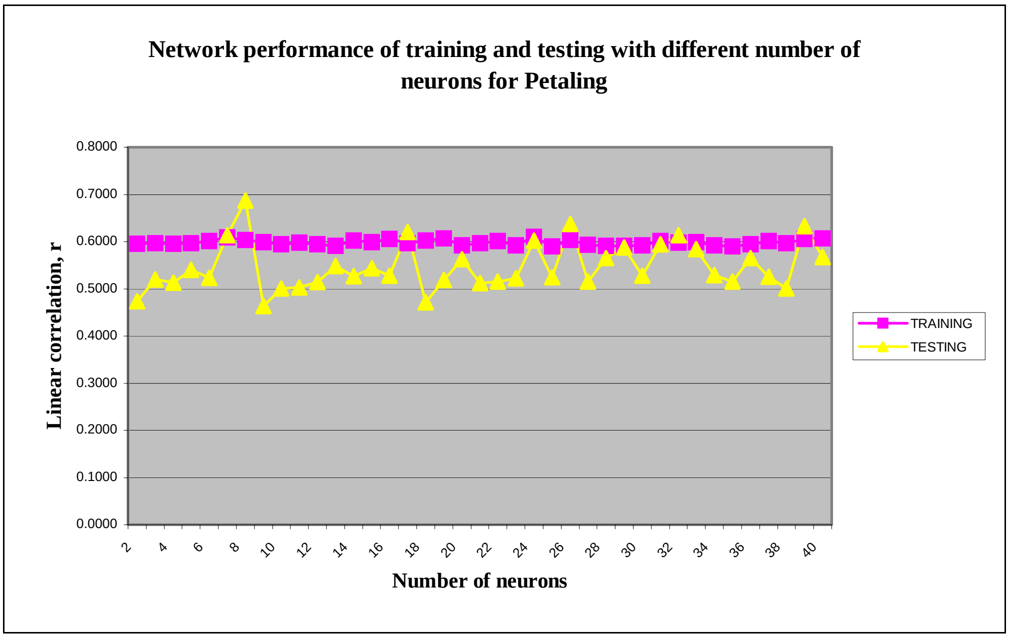

FIGURE 2 Network performance of training and testing with different number of neurons for Petaling

FIGURE 3 Network performance of training and testing with different number of neurons for Gombak

From table 2, the value of MAPE shows that Neural Network can forecast demand on low cost housing in Petaling very good with MAPE of 6.42 and 4.32.

Table 3 shows that performance of forecasted demand on low cost housing in Gombak is very good in September 1999 with MAPE value of 8.78 and good in October 1999 with MAPE value 11.83.

Discussion

After several round of trial and error process, the optimum training and testing data for Petaling are 40 and 9 and Gombak are 46 and 3. The best Neural Network to forecast demand on low cost housing in Petaling is 2-8-1and 3-11-1 for Gombak using 0.7 learning rate and 0.4 momentum rate with r value of 0.6796 and 0.8162 respectively. The results of forecasted demand on low cost housing in the districts show that NN can forecast demand on low cost housing very good in Petaling with MAPE of 6.42 in September 1999 and 4.32 in October 1999 while in Gombak, the results is very good in September 1999 with MAPE of 8.78 and good in October 1999 with MAPE of 11.83.

Conclusions

In conclusion, an ANN can forecast demand on low cost housing in district of Petaling very good and good in the district of Gombak. The computerized model can be used to plan the construction of low cost houaing. It also can save cost,time, manpower and paper wastages.

TABLE 2 Forecasted demand on low cost housing for September and October 1999 for Petaling

Time series	Actual	Forecasted	MAPE
September 1999	327	348	6.42
October 1999	324	338	4.32

TABLE 3 Forecasted demand on low cost housing for September and October 1999 for Gombak

Time series	Actual	Forecasted	MAPE
September 1999	330	301	8.78
October 1999	321	283	11.83

References

Chander, R. (1976). "Housing Needs versus Effective Demand in Malaysia." Department of Statistic of Malaysia.

Dutta, S. and Shekhar, S. (1998). "Bond Rating: A Non-Conservative Application of Neural Networks." *Proceedings of the IEEE International Conference on Neural Networks*, p. II:443-450.

Ghani Salleh and Lee, Lik Meng (1997). "Low-cost housing in Malaysia." Kuala Lumpur: Utusan Publications & Distribution Sdn. Bhd.

Hua, B.H. (1998). "Forecasting Residential Construction Demand in Singapore: A Comparative Study of The Accuracy of TimeSeries, Regression and Artificial Neural Network Techniques." *Engineering, Construction and Architechture Management*, 261-275.

Kamal Salih (1976). "Socio-economic Consideration in the Planning, Design and Implementation of Human Settlements and Housing For Lower Income Groups in Malaysia." *The Conference of Human Settlements for the Rakyat in the Lower Income Group*, Paper No. 5.

Karunanithi, N., et. al. (1994). "Neural Networks For River Flow Prediction." *Journal of Computing in Civil Engineering*, Vol.8, No. 2, 201-220.

Gergely, M., Somogyi, Sz., and Kohlh⎾b, R. (1994). "Computerized Property Prediction and Process Planning in Heat Treatment of Steels." *Materials Science Forum*, Trans. Tech. Publications, Switzerland, Vols. 163-165, 657-666.

Yang and Packer (1997). "Applying Artificial Neural Networks to UK Construction Demand

Forecasting (Private Sector)." *COBRA '97*, p. 1-16.

Yahya, K. and M. Z. Abd Majid, (2002). "Forecasting of Low Cost Housing Demand In Urban Areas- Artificial Neural Netwrk And ARIMA Model Approach." *2nd International Conference on Systems Thinking in Management*, 14-19.

Zainun, N.Y. and M. Z. Abd Majid, (2002). "Techniques To Developed Needs Model On Housing For Low Income Group: A Literature And Malaysian Experience." *2nd International Conference on Systems Thinking in Management*, 28-37.

Zainun, N.Y. and M.Z. Abd Majid (2003). "Evaluation On Various Forecasting Models Using Artificial Neural Networks (ANN)" *2nd International Conference on Innovation in Architecture, engineering and construction (AEC),* UK.

Zainun, N.Y. and Muhd Zaimi Abd. Majid (2002). "An Overview of Forecasting Model Using Artificial Neural Networks (ANN)." *Research Seminar on Materials and Construction,* Johor Bahru, Malaysia, p. 210-217.