Housing price prediction using neural networks

Conference Paper · August 2016

DOI: 10.1109/FSKD.2016.7603227

Lipo Wang
Nanyang Technological University

Y.L. Wang
Taiyuan University of Technology

Qing Chang

Predicting Public Housing Prices Using Delayed Neural Networks

Lipo Wang, Fung Foong Chan, Yaoli Wang, and Qing Chang

***Abstract*—This study uses delayed neural network models to predict public housing prices in Singapore. The delayed neural networks are used to estimate the trend of the resale price index (RPI) of Singapore housing from the Singapore Housing Development Board (HDB), with nine independent economic and demographic variables. The results show that the delayed neural network model is able to produce a good fit and predictions.**

***Index Terms*—Artificial neural network (ANN), public housing prices, estimate, Singapore**

I. Introduction

Public housing in Singapore is important to the residents, as more than 80% of the nation's population live in the Housing and Development Board (HDB) flats nowadays. According to the Department of Statistics of Singapore, the population density of Singapore is one of the highest in the world.

Against this background, it could be beneficial to have a reliable prediction of the public housing prices. Based on the pricing predictions, the authority can better develop sustainable public housing plans and forestall any housing bubbles. With such predictions, investors could maximize profits gained from the performed transactions, whereas potential buyers will acquire the basic understanding of the latest pricing trends, enabling them to make informed decisions and reducing the risk of loss.

Traditionally, various approaches have been introduced to forecast the housing prices. Some of the most widely used methods include the sales comparison grid and, more recently, the hedonic pricing models derived from multiple regression analysis. Nonetheless, both of these methods have their respective shortcomings. According to Worzala et al. [1], the sales comparison grid is criticized as inaccurate due to the difficulty in obtaining reliable data. Hedonic pricing, on the other hand, not only is unable to effectively capture the multi co-linearity and non-linearity among the variables, but also involves statistical assumptions in the samples. Therefore, the results generated might be unfavorable.

Artificial neural networks (ANN) are recognized for their learning and generalization abilities [2]. ANNs are able to approximate the mapping of arbitrary nonlinear variables [3-13]. Today, ANNs are used not only for housing prices estimation (e.g., [14-33]), but also in other fields, such as stock price index estimation (e.g., [34-42]) and many types of time series prediction (e.g., [43-48]).

Numerous studies have been conducted to assess the effectiveness of ANNs in housing price predictions. In this context, ANN has been applied for more than 20 years (e.g., [14-33]).

The main objective of this study is to examine the accuracy of ANNs in predicting Singapore's public housing prices. In order to fulfill this objective, we will attempt to identify the possible factors influencing the public housing prices.

In Section 2, a dynamic ANN is used for forecasting, as the time factor is now taken into consideration. It is undeniable that the public housing price varies over time. Therefore, time-series data is used as the input and output variables. Besides that, it is assumed that economic variables will play an important role in determining the price trend of public housing. As a result, several economic and demographic factors are chosen as the inputs.

II. Predicting Public Housing Prices Using Delayed Neural Networks

The housing prices could be affected by the characteristics of a particular housing unit, as well as other economic and demographic factors [49-54].

There are a number factors to be taken in consideration when evaluating the potential selling price of a HDB unit, of which the following six are intrinsic to each unit:

- Number of rooms
- Floor / Level
- Area of floor space
- Duration of elapsed lease ("Age" of the HDB unit)
- Distance to the nearest school
- Distance to the nearest subway (MRT) station

In this study, ten Singapore economic variables are initially considered as exogenous variables, which may exhibit a strong correlation with the housing prices. Before training the ANN, these variables are to undergo a correlation test to determine the most suitable variables for the input vectors. All the ten variables are listed as follows:

- Singapore Real Gross Domestic Product (GDP)
- Population
- Unemployment Rate
- Average Monthly Wages
- Labour Cost
- Straits Times Index (STI) for Singapore stock market
- Prime Lending Rate
- Interbank Rate
- Singapore Consumer Spending
- Singapore Consumer Price Index (CPI)

There are several general steps when developing and implementing the ANN model, as listed below:

Lipo Wang and Fung Foong Chan are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (email: elpwang@ntu.edu.sg)

Yaoli Wang (corresponding author) and Qing Chang (corresponding author) are with College of Information Engineering, Taiyuan University of Technology, Taiyuan, China.

978-1-5090-2597-8/16/$31.00 ©2016 IEEE

1. Data Collection and Preprocessing
2. Network Creation and Configuration
3. Network Training, Validation and Testing
4. Result Analysis

This study consists of two parts, i.e., (1) MATLAB Neural Network Fitting Tool and (2) the MATLAB Neural Network Time Series Tool will be deployed.

The data used for part 1 is retrieved from the HDB and the PropertyGuru websites [49-54]. The statistics of the HDB resale prices are collected for the 26 HDB estates in Singapore, from December 2012 to February 2014.

There are six input (independent) variables to be examined, as indicated earlier. The targeted output variable is the resale price of the HDB unit (in ten thousands of Singapore Dollar, i.e., S$ ’0,000).

Meanwhile, the data used in part 2 is different from that of part 1. In part 2, the independent variables used are listed earlier, while the HDB Resale Price Index (RPI) is chosen as the dependent variable. RPI acts as the single indicator that reflects the movements of public housing prices.

For all the variables, quarterly time-series data from 1990 Quarter 1 to 2013 Quarter 4 are used for the ANN model training, validation and testing. There are a total of 96 time steps for each of the time-series data. These data are collected from HDB and Trading Economy website (online resources) [49-54].

In this analysis, the feed-forward ANN models constructed are of different architecture (in terms of the number of hidden layer and neurons), and different distribution ratio of data samples are used for training, validation and testing purposes. The default training algorithm to be used is the Levenberg-Marquardt algorithm. The two main performance indicators are the Mean Squared Error (MSE) and Regression Value (R-value). Each ANN model is to be trained for several times, and the best-performed ANN is chosen based on the lowest MSE value generated.

Table 1. Exprimental Results

Group	ANN Architecture	Data Sample Distribution Ratio	Average MSE	Overall R-value
1	6-8-1	60:20:20	3.882	0.9554
2	6-8-1	60:20:20	39.90	0.9062
3	6-10-1	60:20:20	21.46	0.9209
4	6-15-1	70:15:15	13.03	0.9579

From Table 1, it can be observed that the results produced by the ANN model are acceptable. For the cases of best-performed ANNs, the R-value lies between 0.9062 and 0.9579. This shows that there are strong correlations between the generated outputs and the targeted outputs. The ANN is able to formulate a good fit between those independent variables with the corresponding dependent variable.

In Figure 1, it can be seen that the ANN is able to produce accurate forecasts for most of the time steps. However, significant errors occurred around the 10th, 20th and 30th time steps. Generally, the predictions are reliable, since most of the error correlations are within the 95% confidence limit (except for the one at zero lag). The R-values for training, validation and testing are 0.997 and above. The low MSE values imply that the ANN model is able to produce predictions close to the actual output targets.

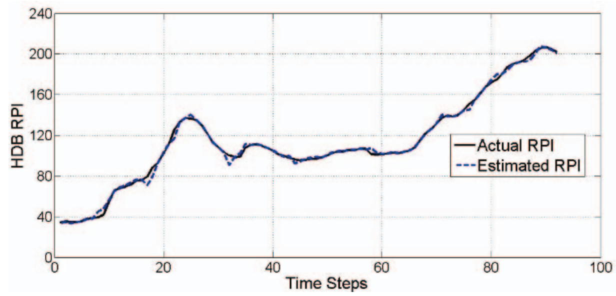

Fig. 1. Plot of RPI against Time Steps

III. Conclusion

The study aims to evaluate the effectiveness of artificial neural network in predicting the Singapore public housing prices. The analysis has shown that the neural network is able to provide estimations that correlate well with the real housing market situation.

For static ANNs, the neural network model is able to map the non-linear relationship between resale price and those housing characteristics that influence the housing price. The R-value for all the best-performed ANN models is higher than 0.9. Therefore, the fitting between independent variables and dependent variables is reasonably good. The performance of the static ANN can be improved after making changes to the training procedure.

For dynamic ANNs, the accuracy of the predictions is high, because the values of predicted RPI are close to the actual target. It has the ability to deduce and generalize the relationship between independent input vectors and the housing price index. The price index movement can therefore be estimated, with a relatively small error.

However, ANN has its limitations. One main problem is the inconsistency in the results. During the training process, the ANN model is capable of self-learning and adjusting the weights accordingly to minimize the error. The initial conditions of the network are totally different for each training, and there are neither formulae nor rules to be followed. As a result, no conclusion can be made on whether the obtained ANN architecture and the results produced are optimal. One may conduct n-fold cross validation or run multiple times with different initial weights.

To summarize, ANN is a useful tool in housing prices prediction and other financial applications. Nonetheless, users must also be aware of its underlying weaknesses, and caution is important when using ANN models for financial forecasting. In future studies, we shall use more rigorous techniques to select input features (e.g., [23-25]) and test other predictive models (e.g., [26-28])

References

[1]. E. Worzala, M. Lenk, and A. Silva, "An Exploration of Neural Networks and Its Application to Real Estate Valuation," *The Journal of Real Estate Research*, vol.10, no.2, pp.185-201, 1995.

[2]. X.J. Ge, G. Runeson, and K.C. Lam, "Forecasting Hong Kong Housing Prices: An Artificial Neural Network Approach," *International Conference of Methodologies in Housing Research*, Stockholm, Sweden, September 22 - September 24, 2003.

[3]. P. Zhang, W. Ma, and T. Zhang, "Application of Artificial Neural Network to Predict Real Estate Investment in Qingdao," In Y. Zhang (ed.), *Future Communication, Computing, Control and Management*, vol.1, pp.213-219, Springer-Verlag Berlin Heidelberg, 2012.

[4]. F. Chu and L. P. Wang, "Applications of support vector machines to cancer classification with microarray data," *International Journal of Neural Systems*, vol.15, no.6, pp.475-484, 2005.

[5]. N. Zhou and L. P. Wang, "Effective selection of informative SNPs and classification on the HapMap genotype data," *BMC Bioinformatics,* 8:484, 2007.

[6]. L. P. Wang, "Learning and retrieving spatio-temporal sequences with any static associative neural network," *IEEE Trans. Circuit and Systems-II: Analog and Digital Signal Processing,* vol. 45, no.6, pp. 729-738, June, 1998.

[7]. L. P. Wang, F. Chu, and W. Xie, "Accurate cancer classification using expressions of very few genes," *IEEE-ACM Trans. Computational Biology and Bioinformatics*, vol.4, no.1, pp. 40-53, Jan.-March, 2007.

[8]. B. Liu, C.R. Wan, and L. P. Wang, "An efficient semi-unsupervised gene selection method via spectral biclustering", *IEEE Trans. Nano-Bioscience*, vol.5, no.2, pp.110-114, June, 2006.

[9]. L. P. Wang and X. J. Fu, *Data Mining with Computational Intelligence*, Berlin/Heidelberg: Springer-Verlag, 2005.

[10]. X.J. Fu and L. P. Wang, "Data dimensionality reduction with application to simplifying RBF network structure and improving classification performance", *IEEE Trans. Systems, Man, Cybern, Part B-Cybernetics,* vol.33, no.3, pp. 399-409, 2003.

[11]. F. Chu, Wei Xie, and L.P. Wang, "Gene selection and cancer classification using a fuzzy neural network", *Proceedings of the North-American Fuzzy Information Processing Conference (NAFIPS 2004),* vol.2, pp.555-559, 2004.

[12]. L. P. Wang, Nina Zhou, and Feng Chu, "A general wrapper approach to selection of class-dependent features," *IEEE Trans. Neural Networks*, vol.19, no.7, pp.1267-1278, 2008.

[13]. Y. Frayman and L. P. Wang, "Data mining using dynamically constructed recurrent fuzzy neural networks," *Research and Development in Knowledge Discovery and Data Mining*, vol. 1394, PAKDD'98 (Regular Paper), pp. 122-131, 1998.

[14]. V. Limsombunchai, "House Price Prediction: Hedonic Price Model vs. Artificial Neural Network," *2004 New Zealand Agriculture and Resource Economics Society (NZARES) Conference*, Blenheim, New Zealand, June 25 – June 26, 2004.

[15]. V. Limsombunchai, C. Gan, M. Lee, "House Price Prediction: Hedonic Price Model vs. Artificial Neural Network," *American Journal of Applied Sciences*, vol.1, pp.193-201, 2004.

[16]. A. Khalafallah, "Neural network based model for predicting housing market performance," *Tsinghua Science and Technology*, vol.13, no.S1, pp.325 – 328, 2008.

[17]. N. Nguyen and A. Cripps, "Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networks," *Journal of Real Estate Research*, vol.22, no.3, pp.313-336, 2001.

[18]. Y. El Hamzaoui and J.A. Hernandez-Perez, "Application of Artificial Neural Net-works to Predict the Selling Price in the Real Estate Valuation Process," *2011 10th Mexican International Conference on Artificial Intelligence (MICAI)*, pp.175 – 181, 2011.

[19]. P. Lai, "Analysis of the Mass Appraisal Model by Using Artificial Neural Network in Kaoshiung City," *Journal of Modern Accounting and Auditing*, 7(10), pp.1080-1089, 2011.

[20]. K.C. Lam, C. Y. Yu, and K.Y. Lam, "An Artificial Neural Network and Entropy Model for Residential Property Price Forecasting in Hong Kong," Journal of Property Research, 25(4), pp. 321-342, 2008.

[21]. A. Evans, H. James, and A. Collins, "Artificial Neural Networks: An Application to Residential Valuation in the UK," Journal of Real Estate Valuation and Investment, 11(2), pp.195-204, 1991.

[22]. A. Do and G. Grudnitski, "A Neural Network Approach to Residential Property Appraisal," The Real Estate Appraisal, 58(3), pp.38-45, 1992.

[23]. D.P.H. Tay and D.K.K. Ho, Artificial Intelligence and the Mass Appraisal of Resi-dential Apartments. Journal of Property Valuation and Investment, 10(2), pp.525-540, 1992.

[24]. P.O. Eriki and R.I. Udegbunam, "Application of Neural Network in Evaluating Prices of Housing Units in Nigeria: A Preliminary Investigation," Journal of Artificial Intelligence, 1(1), pp. 21-27, 2008.

[25]. H. Selim, "Determinants of House Prices in Turkey: Hedonic Regression versus Artificial Neural Network," Expert Systems with Applications, 36, pp.2843-2852, 2009.

[26]. MathWorks, Inc. Neural Network Time Series Prediction and Modeling. http://www.mathworks.com/help/nnet/gs/neural-network-time-series-prediction-and-modeling.html. Retrieved on March 25, 2014.

[27]. P. Rossini, "Application of Artificial Neural Networks to the Valuation of Residential Property," The Third Annual Pacific-Rim Real Estate Society Conference, Palmerston North, New Zealand, January 20-22, 1997.

[28]. I.D. Wilson, S.D. Paris, J.A. Ware, and D.H. Jenkins, "Residential Property Price Time Series Forecasting with Neural Networks," In A. Macintosh et al. (eds.), Applications and Innovations in Intelligent Systems IX (pp.17-28). London: Springer-Verlag London, 2002.
[29]. K.C. Wong, A.T.P. So, and Y.C. Hung, "Neural Network vs. Hedonic Price Model: Appraisal of High-Density Condominiums," In K. Wang et al. (eds.), Real Estate Valuation Theory (pp.181-198). New York, NY: Springer Science + Business Media, 2002.
[30]. T.-C. Wong, A. Yap, "From universal public housing to meeting the increasing aspiration for private housing in Singapore," Habitat International, vol. 27, pp. 361–380, 2003.
[31]. Y. Yan, W. Xu, H. Bu, Y. Song, W. Zhang, H. Yuan, S.-Y. Wang, "Method for Housing Price Forecasting based on TEI@I Methodology," Systems Engineering - Theory & Practice, pp.1-9, 2007.
[32]. J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, New York: Springer Science, 2002.
[33]. B. Sun, "A Study about the Prediction of University Library Lending Based on Multiple Regression Analysis," In B. Sun, Advances in Automation and Robotics, vol.1, pp.525-532, Springer Berlin Heidelberg, 2012.
[34]. A. Aussem, J. Campbell, and F. Murtagh, "Wavelet-based feature extraction and decomposition strategies for financial forecasting," J. Comput. Intell. Finance, vol. 6, no. 2, pp. 5–12, 1998.
[35]. L. P. Wang and S. Gupta, "Neural Networks and Wavelet De-Noising for Stock Trading and Prediction," in Time Series Analysis, Modeling and Applications, W. Pedrycz and S.-M. Chen, Eds. Springer Berlin Heidelberg, 2013, pp. 229–247.
[36]. S. Gupta and L. P. Wang, "Stock Forecasting with Feedforward Neural Networks and Gradual Data Sub-Sampling," Australian Journal of Intelligent Information Processing Systems, vol.11, pp.14-17, 2010.
[37]. G. Dong, K. Fataliyev, and L. P. Wang, "One-step and multi-step ahead stock prediction using backpropagation neural networks," in Communications and Signal Processing (ICICS) 2013 9th International Conference on Information, 2013, pp. 1–5.
[38]. Y. Fang, K. Fataliyev, L.P. Wang, X.J. Fu and Yaoli Wang, "Improving the genetic-algorithm-optimized wavelet neural network approach to stock market prediction," 2014 International Joint Conference on Neural Networks (IJCNN 2014), pp. 3038-3042.
[39]. T. Zheng, K. Fataliyev, and L. P. Wang, "Wavelet neural networks for stock trading and prediction," SPIE Defense, Security, and Sensing, 29 April - 3 May 2013, Baltimore, USA, vol.8750, 0A (8750-9).
[40]. M. Zhu and L. P. Wang, "Intelligent trading using support vector regression and multilayer perceptrons optimized with genetic algorithms," 2010 International Joint Conference on Neural Networks (IJCNN), 2010, pp. 1–5.
[41]. Y.Q. He, K. Fataliyev, and L. P. Wang, "Feature selection for stock market analysis," the 20th International Conference on Neural Information Processing (ICONIP2013), Daegu, Korea, 3-10 November 2013, Invited Paper, Part II, LNCS 8227, pp. 737–744, 2013.
[42]. H. Zhou and Y. Wei, "Stocks market modeling and forecasting based on HGA and wavelet neural networks," in 2010 Sixth International Conference on Natural Computation (ICNC), 2010, vol. 2, pp. 620–625.
[43]. A. B. Geva, "ScaleNet-multiscale neural-network architecture for time series prediction," IEEE Trans. Neural Netw., vol. 9, no. 6, pp. 1471–1482, Nov. 1998.
[44]. K. K. Teo, L. P. Wang, and Z. Lin, "Wavelet Packet Multi-layer Perceptron for Chaotic Time Series Prediction: Effects of Weight Initialization," in Computational Science - ICCS 2001, V. N. Alexandrov, J. J. Dongarra, B. A. Juliano, R. S. Renner, and C. J. K. Tan, Eds. Springer Berlin Heidelberg, 2001, pp. 310–317.
[45]. L. P. Wang, K.K. Teo, and Z.P. Lin, "Predicting time series with wavelet packet neural networks", 2001 IEEE International Joint Conference on Neural Networks (IJCNN 2001), pp.1593-1597, 2001.
[46]. G. Zhang, B.E. Patuwo, and M.Y. Hu, "Forecasting with Artificial Neural Net-works: The State of the Art," International Journal of Forecasting, 14, pp.185-201, 1998.
[47]. A. Katchova, "Time Series ARIMA Models", https://drive.google.com/file/d/0BwogTI8d6EEiaDJCRXd0dmU1ZDA/edit?pli=1, 2013.
[48]. M. Majidpour, C. Qiu, P. Chu, R. Gadh, H.R. Pota, "Fast Prediction for Sparse Time Series: Demand Forecast of EV Charging Stations for Cell Phone Applications," IEEE Transactions on Industrial Informatics, vol.11, pp.242-250, 2015.
[49]. PropertyGuru, http://www.propertyguru.com.sg/singapore-property-listing/hdb. Retrieved on March 10, 2014.
[50]. Singapore Housing & Development Board (HDB). http://services2.hdb.gov.sg/webapp/BB33RTIS/BB33PReslTrans.jsp. Retrieved on March 10, 2014.
[51]. Government of Singapore, Department of Statistics, http://www.singstat.gov.sg/statistics/latest_data.html, 31 Oct, 2014.
[52]. Singapore Department of Statistics, Singapore in Figures 2013 [PDF Document]. http://www.singstat.gov.sg/publications/publications_and_papers/reference/sif2013.ppd. Retrieved on March 20, 2014.
[53]. STProperty Condo Search, http://www.stproperty.sg/condominium-directory/top-from-1970/top-to-2014/sort-top-asc/page1/box-1
[54]. Trading Economics, http://www.tradingeconomics.com/ singapore/indicators. Retrieved on March 18, 2014.