applied sciences

MDPI

Article

Using Machine Learning Models and Actual Transaction Data for Predicting Real Estate Prices

Ping-Feng Pai * and Wen-Chang Wang

Department of Information Management, National Chi Nan University, 1 University Rd., Puli, Nantou 54561, Taiwan; s107213502@mail1.ncnu.edu.tw

* Correspondence: paipf@ncnu.edu.tw

Received: 22 July 2020; Accepted: 20 August 2020; Published: 23 August 2020

Abstract: Real estate price prediction is crucial for the establishment of real estate policies and can help real estate owners and agents make informative decisions. The aim of this study is to employ actual transaction data and machine learning models to predict prices of real estate. The actual transaction data contain attributes and transaction prices of real estate that respectively serve as independent variables and dependent variables for machine learning models. The study employed four machine learning models-namely, least squares support vector regression (LSSVR), classification and regression tree (CART), general regression neural networks (GRNN), and backpropagation neural networks (BPNN), to forecast real estate prices. In addition, genetic algorithms were used to select parameters of machine learning models. Numerical results indicated that the least squares support vector regression outperforms the other three machine learning models in terms of forecasting accuracy. Furthermore, forecasting results generated by the least squares support vector regression are superior to previous related studies of real estate price prediction in terms of the average absolute percentage error. Thus, the machine learning-based model is a substantial and feasible way to forecast real estate prices, and the least squares support vector regression can provide relatively competitive and satisfactory results.

Keywords: real estate prices; machine learning; predict

1. Introduction

The real estate market is one of the most crucial components of any national economy. Hence, observations of the real estate market and accurate predictions of real estate prices are helpful for real estate buyers and sellers as well as economic specialists. However, real estate forecasting is a complicated and difficult task owing to many direct and indirect factors that inevitably influence the accuracy of predictions. In general, factors influencing real estate prices could be quantitative or qualitative [1]. The quantitative factors possibly include macroeconomic factors [2], business cycles [3], and real estate attributes [4]. The macroeconomic factors contain unemployment rates, share index, current account of a country, industrial production, and gross domestic product [2]. Attributes of real estate, for example, includes past sale prices, land area, years of constructions, floor space, surface area, number of floors and building conditions [1,4]. The qualitative factors refer to subject preferences of decision makers, such as views [5], building styles, and living environment [1]. However, some difficulties arise in data collection for qualitative factors. For qualitative factors, sometimes these data are suffering from lack of measurements. Data of qualitative factors sometimes are suffering from lack of measurements. Thus, qualitative factors are hard to measure [1]. Therefore, this study did not take qualitative factors influencing real estate prices into considerations and used quantitative data gathered from actual transaction data recording details of real estate transaction data in Taiwan. Four machine learning models were used to forecast real estate prices accordingly. The rest of this

study is organized as follows. Section 2 presents literature review of real estate prediction. Section 3 illustrates methods used in this study. Section 4 introduces the proposed real estate appraising system. The numerical results are depicted in Section 5. Section 6 provides conclusions.

2. The Literature Review of Real Estate Price Predictions

Some studies of real estate prices predictions are presented as follows. Singh et al. [6] employed the concept of big data to predict housing sale data in Iowa, using three models to forecast house sale prices: linear regression, random forest, and gradient boosting. The numerical results indicated that the gradient boosting model outperforms the other forecasting models in terms of forecasting accuracy. Segnon et al. [7] presented a logistic smooth transition autoregressive fractionally integrated process to predict housing price volatility in the U.S.A. and analyzed complicated statistical models based on assumptions of the variance process. The numerical results revealed that the Markov-switching multifractal and fractionally integrated generalized autoregressive conditional heteroscedastic models provide satisfied forecasting accuracy.

Kang [8] developed a news article-based forecast model to predict Jeonse prices in South Korea. The Internet search intensity of keywords from news was treated as the independent variable. The numerical results showed that the designed models obtain more accurate results than time series techniques. Giudice et al. [9] used genetic algorithms to forecast real estate rental prices when geographic locations and four real estate attributes are considered. The multivariate regression technique was performed to forecast the same data. Numerical results indicated that the genetic algorithms are superior to the multivariate regression in term of prediction accuracy. Park and Bae [10] designed a house price prediction system by machine learning approaches to help house sellers or real estate agents with house price evaluations. In their investigation, data were collected from the Multiple Listing Service of the Metropolitan Regional Information Systems. The numerical results indicated that repeated incremental pruning to produce error reduction (RIPPER) algorithm obtains more accurate forecasting results than the other forecasting models.

Bork and Møller [11] employed dynamic model averaging and dynamic model selection to forecast house price growth rates in the 50 states of the U.S.A. The presented forecasting system captures house price growth rates by varying the model and the coefficients' change over time and across locations. Thus, the forecasting results provided by the proposed system are substantial. Plakandaras et al. [12] designed a hybrid model as an early warning system, including ensemble empirical mode decomposition and support vector regression, for predicting sudden house price drops. The proposed forecasting approach generates superior results over other forecasting models in terms of forecasting accuracy. Chen et al. [13] developed a housing price analysis system using information from public websites to analyze the total ratio of the average value and standard deviation of housing prices. They reported that the designed housing price analysis system is a helpful way for obtaining insights into housing prices. Lee et al. [14] employed fuzzy adaptive networks to forecast pre-owned housing prices in Taipei by taking both objective variables and subjective variables into considerations. The empirical results indicated that the fuzzy adaptive networks outperform back-propagation neural networks and the adaptive network fuzzy inference system in terms of forecasting accuracy.

Antipov and Pokryshevskaya [15] used random forest to appraise residential estate of Saint Petersburg, Russia and reported that the random forest approach outperforms the other forecasting methods in prediction accuracy. In addition, the authors claimed that the random forest approach is capable of dealing with missing values and categorical variables. Kontrimas and Verikas [16] presented an ensemble learning system integrating ordinary least squares linear regression and support vector regression to appraise real estate. In addition, weights based on value zones provided by experts of the register center and weights generated by the self-organizing map (SOM) were used by the ensemble learning system. The numerical results revealed that the ensemble learning system with weights generated by SOM outperform the ensemble learning systems with weights provided by

experts. Furthermore, the ensemble systems can reach more accurate forecasting results than the other single forecasting models.

Gupta et al. [17] employed time series models with or without the information content of 10 or 120 additional quarterly macroeconomic series to forecast the U.S. real house price index. Their study concluded that the utilization of fundamental economic variables could increase the forecasting accuracy and especially be effective for the 10 fundamental economic variables in the dynamic stochastic general equilibrium model. Kusan et al. [18] developed a grading fuzzy logic model to forecast house selling prices. Many housing attributes, such as public transportations systems and environmental factors, served as inputs of the fuzzy logic system. The numerical results indicated that the proposed fuzzy logic system captures the patterns of house selling prices and lead to satisfied forecasting accuracy. The collection of actual transaction data of real estate has been performed since August 2012 by the Ministry of the Interior of Taiwan. Thus, the motivation of this study is to examine the performance of machine learning models with actual transaction data in forecasting real estate prices.

3. Methods

3.1. Least Squares Support Vector Regression

The support vector machines [19,20] technique was originally designed for classification problems. For dealing with regression problems, support vector regression [21–23] was proposed and has become a popular alternative for estimating linear or non-linear prediction problems. However, owing to solving quadratic functions in the procedure of support vector regression, the computation burden is quite challenging. Thus, the least squares support vector regression [24] was designed to decrease the computation load by converting a quadratic programming problem into a linear problem. For an input-output dataset $\{X_i, Y_i\}$, $i = 1 \ldots N$, the LSSVR model can be illustrated as Equation (1) [24].

$$\text{Min}: \ f(W,\tau) = \frac{1}{2}W^T W + \frac{1}{2}\Omega\sum_{i=1}^{N}\tau_i^2 \tag{1}$$

Subject to

$$y_i = W^T{\cdot}\psi(x_i) + \delta + \tau_i, i = 1,\ldots,N$$

where W represents the weighted vector, Ω denotes the penalty factors controlling the trade-off between the optimization of approximation error and flatness of the approximated function, τ_i is the ith error vector, $\psi(x_i)$ is the nonlinear mapping function transferring the original input space into a high dimension input space, δ indicates the bias parameter. Using the Lagrange function, the Lagrange form of Equation (1) is depicted as Equation (2)

$$L(W,\delta,\alpha,\ \tau) = f(W,\tau) + \sum_{i=1}^{N}\alpha_i\Big(y_i - W^T{\cdot}\psi(x_i) - \delta - \tau_i\Big) \tag{2}$$

where α_i represent Lagrange multipliers.

Applying the Karush–Kuhn–Tucker conditions [25–27] and setting derivatives with respect to four variables, $W, \delta, \alpha,$ and τ, equal to zero, Equations (3)–(6) can be obtained.

$$\frac{\partial L(W,\delta,\alpha,\ \tau)}{\partial W} = 0,\ W = \sum_{i=1}^{N}\delta_i\psi(x_i) \tag{3}$$

$$\frac{\partial L(W,\delta,\alpha,\ \tau)}{\partial \delta}\ ,\ \sum_{i=1}^{N}\alpha_i = 0 \tag{4}$$

$$\frac{\partial L(W,\delta,\alpha,\ \tau)}{\partial \tau}\ ,\ \alpha_{\mathrm{i}} = \Omega\tau_i \tag{5}$$

$$\frac{\partial L(W,\delta,\alpha,\ \tau)}{\partial \delta}\ ,\ W^T\psi(x_i) + \delta + \tau_i - y_i = 0 \tag{6}$$

Then, by the least squares method, the LSSVR model can be reformed as Equation (7):

$$y(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i) + \delta \tag{7}$$

$$K(x, x_i)\ =\ \psi(x)^T \cdot \psi(x_i)^T \tag{8}$$

where $K(x, x_i)$ represents the kernel function satisfying the Mercer's principle [28]. The radial basis function with a variance σ^2 expressed by Equation (9) is specified here as a kernel function.

$$K(x, x_i) = exp(-\frac{\|x - x_i\|^2}{2\sigma^2}) \tag{9}$$

3.2. Classification and Regression Trees

Proposed by Breiman et al. [29], the classification and regression tree is one of the most popular techniques in dealing with classification or regression problems. The Gini measurement and least-squared deviation measurement are performed by CART models for categorical and numerical problems correspondingly [29,30]. Let the p^{th} sample be illustrated as $\left(I_{p,1}, I_{p,2}, \ldots\ldots I_{p,n} \ldots . O_p\right)$, where $I_{p,n}$ is the value of the p^{th} sample with n features, and O_p is the corresponding output value of the sample. For a regression problem of CART, the minimization of the least-squared deviation measure of impurity [29] represented by Equation (10) serves as a decision to determine the split-up of trees into branches.

$$\frac{1}{N}\sum_{\mathrm{V}\in \mathrm{U}_r}\left(O_p - \overline{O}_r\right)^2 + \frac{1}{N}\sum_{\mathrm{V}\in \mathrm{U}_l}\left(O_p - \overline{O}_l\right)^2 \tag{10}$$

where N is the total number of training samples, U_r and U_l are training data sets directing to the right child node and the left child node correspondingly. $\overline{O}_r$ and $\overline{O}_l$ are mean output values of the right node and the left node respectively.

3.3. General Regression Neural Networks

Based on the Parzen window non-parametric estimation [31] of a probability density function, the general regression neural network [32] is a probabilistic neural network that is able to cope with linear or non-linear forecasting problems with continuous outcome values. Suppose a probability density function $f(I, O)$ is associated with input vectors I and output vectors O. The regression of O on i is illustrated as Equation

$$E(o|I) = \frac{\int_{-\infty}^{\infty} o f(I,o) d_o}{\int_{-\infty}^{\infty} f(I,o) d_o} \tag{11}$$

where I is the expected value of i.

Furthermore, the process of conducting a general regression neural network model can be treated as dealing with the kernel regression problems expressed as Equation (12), where (I_p,O_p) is a data pair of (I,O), and σ is the smoothing parameter.

$$O\ (I)\ =\ \frac{\sum_{p=1}^{q} O_p \exp\left[-\ \frac{(I-\ I_p)^T (I\ -\ I_p)}{2\sigma^2}\right]}{\sum_{p=1}^{q} \exp\left[-\ \frac{(I-\ I_p)^T (I\ -\ I_p)}{2\sigma^2}\right]} \tag{12}$$

where (I_p, O_p) is a data pair of (I, O), σ is the smoothing parameter.

3.4. Backpropagation Neural Networks

The backpropagation learning algorithms [33] make up a powerful and popular method to train multilayer perceptron neural networks. The backpropagation neural network containing one or more hidden layers delivers data patterns forward from the input layer through hidden layers to the output layer and generates output values. Feedback errors represented by the difference between the actual output values and the output values of the networks are then sent backward from the direction of the output layer to the input layer. During the process errors' propagation, a chain rule is applied to obtain the updated weights between neuros. By passing data forward and transmitting errors backward iteratively, the training error value decreases. For simplicity and generality, three layers MLP is illustrated for addressing the learning procedures of backpropagation neural networks [34]. Suppose the input data set is represented as I, the *h-th* hidden neuro obtains an input depicted as Equation (13):

$$HNI_h = f\left(\sum_{p=1}^{l} U_{hp} I_p + B_h\right) = f(NET_h), p = 1 \ldots l, h = 1 \ldots m \tag{13}$$

The $f(\cdot)$ is the activation function, B_h is the bias of the *h-th* hidden neuro, I_p is the *p-th* input value, U_{hp} is the weight between the *p-th* input neuro and the *h-th* hidden neuro.

The output of the *h-th* hidden neuro is represented as Equation (14).

$$HNO_q = \sum_{h=1}^{m} V_{qh} f\left(\sum_{p=1}^{l} U_{hp} I_p + B_h\right), h = 1 \ldots m, q = 1 \ldots n \tag{14}$$

Then, the output of the *q-th* output neuro is represented as Equation (15).

$$Y_q = f\left(\sum_{h=1}^{l} V_{qh} HNI_h + B_q\right), h = 1 \ldots m, q = 1 \ldots n \tag{15}$$

where n is the number of output nodes, Y_q is the output value of the *q-th* output neuro, A_q is the *q-th* actual value.

The training error is expressed as Equation (16).

$$E = \frac{1}{2} \sum_{q=1}^{n} \left(A_q - Y_q\right)^2 \tag{16}$$

Then, by the gradient-descent method and chain rules, updated weights between the output layer and the hidden layer can be obtained and illustrated as Equation (17) and Equation (18).

$$\Delta V_{qh}(t) = \left(\frac{\partial E}{\partial V_{qh}(t)}\right) = -\left(\frac{\partial E}{\partial Y_q}\right)\left(\frac{\partial Y_q}{\partial HNO_q}\right)\left(\frac{\partial HNO_q}{\partial U_{hp}(t)}\right) = \left(A_q - Y_q\right) \acute{f}\left(HNO_q\right)(HNI_h) \tag{17}$$

$$V_{qh}(t+1) = \Omega V_{qh}(t) + \gamma \Delta V_{qh}(t) \tag{18}$$

where $V_{qh}(t+1)$ is the weight connecting hidden neuro h and output neuro q for (t + 1)-*th* epoch, Ω is the momentum, γ denotes the learning rate.

Then, the learning algorithm between the input layer and the hidden layer can be illustrated as Equations (19) and (20).

$$\begin{aligned} \Delta U_{hp}(t) &= -\left(\frac{\partial E}{\partial U_{hp}(t)}\right) = -\left(\frac{\partial E}{\partial HNI_h}\right)\left(\frac{\partial HNI_h}{\partial NET_h}\right)\left(\frac{\partial NET_h}{\partial U_{hp}(t)}\right) \\ &= \sum_{q=1}^{n} \left[\left(A_q - Y_q\right) \acute{f}\left(HNO_q\right) V_{qh}(t)\right] \acute{f}(NET_h)\, I_p \end{aligned} \tag{19}$$

$$U_{hp}(t+1) = \Omega U_{hp}(t) + \gamma \Delta U_{hp}(t) \quad (20)$$

where $U_{hp}(t+1)$ is the weight between input neuro p and hidden neuro h for (t+1)-*th* epoch.

4. Data Collection and the Proposed Real Estate Appraising Framework

4.1. Data Collection

The data were collected from actual transaction price data of Taichung, Taiwan (https://lvr.land.moi.gov.tw/) from 2016/04 to 2019/04. Real estates with buildings served as objectives of this study. Due to blank, unclear and outlier data, data cleansing was conducted before using the data. Blank and unclear data were deleted, and the interquartile range technique was performed to deal with outlier data. After data cleansing, in total 32,215 data observations were used in this study. In addition, the real estate attributes were rearranged, and we finally utilized a total of twenty-three independent variables and one dependent variable to forecast real estate transaction prices. For example, the real estate addresses were transformed into geographical coordinates by the Taiwan Geospatial One-Stop system (https://www.tgos.tw/tgos/Web/Address/TGOS_Address.aspx) provided by the Ministry of the Interior of Taiwan. Three attributes namely "non-urban use land zone", "urban use land zone", and "non-urban land use compilation" were integrated into an attribute of "purpose of land use". Ages of real estate were generated from attributes of transaction dates and completion dates of buildings. Table 1 depicts variables used in this study. Moreover, the Pearson correlation coefficient was used to select independent variables with significant coefficient of correlation. Pearson correlation coefficients of independent variables to the real estate price are listed in Table 2. Independent variables with absolute values of Pearson correlation coefficients larger than 0.1 were left. Totally, eleven independent variables were selected.

Table 1. Depictions of variables.

Types of Variables	Codes of Variables	Descriptions Variables	Data Types
Independent Variables	X_1	City or Township	Categorical
	X_2	With or without parking space	Categorical
	X_3	Longitude	Numerical
	X_4	Latitude	Numerical
	X_5	Transaction area of land	Numerical
	X_6	Purpose of land use	Categorical
	X_7	Ages of buildings	Numerical
	X_8	Transaction amount of property	Numerical
	X_9	Transaction floors	Numerical
	X_{10}	Total floors of buildings	Numerical
	X_{11}	Types of buildings	Categorical
	X_{12}	Use of buildings	Categorical
	X_{13}	Materials of buildings	Categorical
	X_{14}	Total transaction areas of buildings	Numerical
	X_{15}	Number of bedrooms	Numerical
	X_{16}	Number of living rooms	Numerical
	X_{17}	Number of bathrooms	Numerical
	X_{18}	With or without compartments	Categorical
	X_{19}	With or without management committee	Categorical
	X_{20}	Prices per square meter	Numerical
	X_{21}	Types of parking space	Categorical
	X_{22}	Area of parking space	Numerical
	X_{23}	Prices of parking space	Numerical
The Dependent Variable	Y	Transaction prices	Numerical

Table 2. Pearson correlation coefficient of independent variables to the real estate price.

Independent Variables (Pearson Correlation Coefficient)	X_1(0.01158), X_2(0.4097), X_3(−0.043), X_4(−0.0184), X_5(0.3518), X_6(−0.0319), X_7(−0.2897), X_8(0.00465), X_9(0.03199), X_{10}(0.09087), X_{11}(−0.0271), X_{12}(0.0884), X_{13}(−0.0382), X_{14}(0.7923), X_{15}(0.4766), X_{16}(0.4391), X_{17}(0.4024), X_{18}(0.05456), X_{19}(−0.0201), X_{20}(0.5299), X_{21}(0.3871), X_{22}(0.3686), X_{23}(0.1932)

4.2. The Proposed Real Estate Appraising Framework

Figure 1 illustrates the proposed real estate appraising framework in this study. After completing the preparation of data, data were spitted into a training data set and a testing data set with data sizes of eighty percent and twenty percent of the total data correspondingly. Then, a 5-fold cross validation was performed to examine the robustness of machine learning models in forecasting real estate pierces. In this study, genetic algorithms (GA) [35] were used to tune parameters of machine learning models. The average absolute percentage error served as the objective function of genetic algorithms. In this study, parameters of machine learning models were expressed by a chromosome including ten genes in a binary-coded forms, and the population size is twenty. A single point crossover was conducted and the crossover and mutation rates were 0.6 and 0.6, correspondingly. In the parameter tuning procedure, tentative models with different parameters were iteratively generated by genetic algorithms. Parameters generated by genetic algorithms for four machine learning models are specified as follows. Two parameters, penalty factors and widths of Gaussian functions, were selected for the least squares support vector regression models. For the classification and regression tree models, the maximum depth of the tree, the minimum numbers of leaf nodes, and the numbers of predictors were adjusted. The smoothing parameter was determined for general regression neural networks. Learning rates and momentums were altered for backpropagation neural networks.

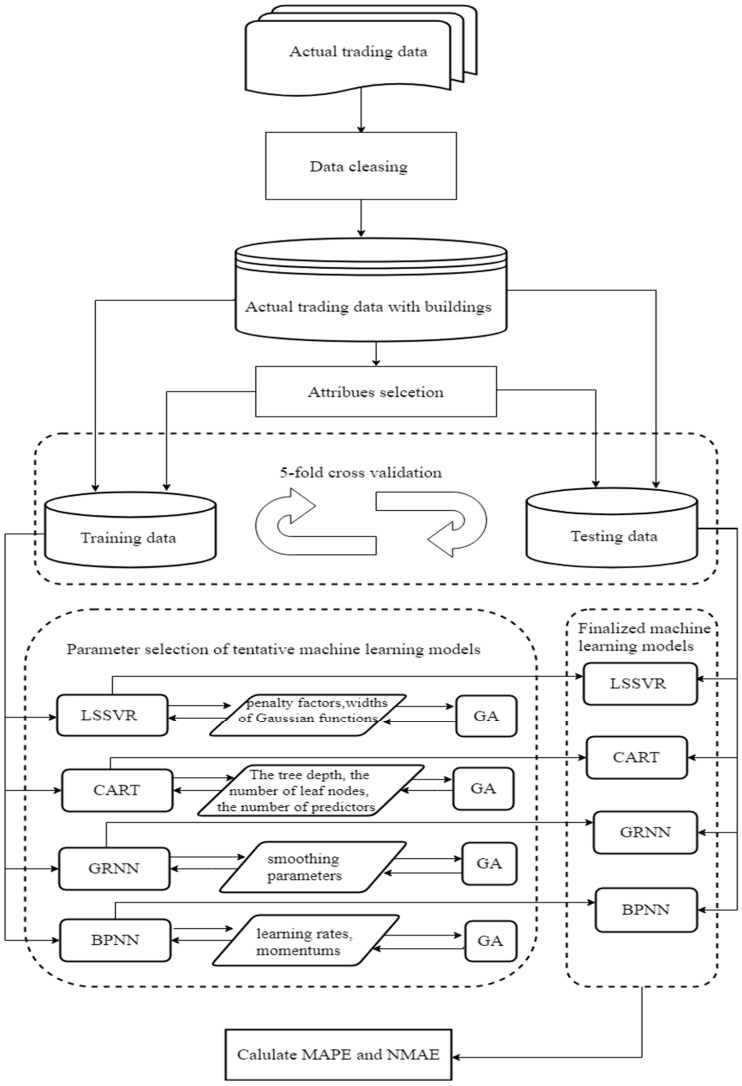

Figure 1. The proposed real estate appraising framework for real estate appraisal.

5. Numerical Results

Two indices, namely the average absolute percentage error (MAPE) and the normalized mean absolute error (NMAE), were employed to evaluate performances of machine learning models in real estate prices forecasting. The mathematical forms of MAPE and NMAE are expressed as follows.

$$\text{MAPE}(\%) = \frac{100}{N}\sum_{t=1}^{N}\left|\frac{R_t - P_t}{R_t}\right| \tag{21}$$

$$\text{NMAE} = \frac{1}{R_h - R_l}\left[\frac{1}{\text{N}}\sum_{t=1}^{N}|R_t - P_t|\right] \tag{22}$$

where N is the amount of forecasting periods, A_t is the real value at period t, and P_t is the predicting value at period t, R_h is the highest actual value, and R_l is the lowest actual value. Tables 3 and 4 show average MAPE and average NMAE values of machines learning models without and with attribute selection correspondingly. The computation results revealed that four machine learning models can result in better results with selected independent variables. According to Lewis [36], MAPE values less than 10 percent are highly accurate predictions; and values between 10 percent and 20 percent are good predictions. Thus, with attribute selection, LSSVR, GRNN and CART are highly accurate predictions and BPNN is a good forecasting in this study. Moreover, real estate prices can be expressed in different currencies or units. To avoid influences of real estate prices in various currencies or units, MAPE is specified when comparing the prediction accuracy of this study with forecasting results of previous studies. This study collected previous related studies which employed MAPE as measurements. Table 5 lists MAPE values of previous studies and the LSSVR models in this study for real estate prices forecasting. Table 5 revealed that the LSSVR models are superior to the other forecasting models of previous studies in terms of MAPE. However, the performance differences may come from many variances of real estate such as countries, cultures, market trends, and economic conditions. Furthermore, buyers' and owners' anticipations are continually varying owning to changes in lifestyles, materials of buildings, environmental legislation, regulations on the rational use of energy [5]. Thus, this study provides a feasible and comparative alternative in forecasting real estate. Forecasting models should be kept adjusted and improved to maintain stability and feasibility over different time periods. In this study, the number of residential properties is 31,397 and the proportion of residential properties is about 97.46 percent of the total data. The other 2.54 percent is commercial real estate. Commercial real estate appraisal uses factors such as spatial autocorrelation which is different from residential real estate appraisal [37–39]. Therefore, the focus of this study is on the residential aspect.

Table 3. Average MAPE and average NMAE values of machines learning models without attribute selection.

Models	LSSVR	CART	GRNN	BPNN
MAPE (%)	1.676	2.2944	22.8936	15.0357
NMAE	4.13×10^{-3}	6.86×10^{-3}	1.07×10^{6}	3.82×10^{-2}

Table 4. Average MAPE and average NMAE values of machines learning models with attribute selection.

Models	LSSVR	CART	GRNN	BPNN
MAPE (%)	0.228	2.278	8.738	14.424
NMAE	8.11×10^{-4}	6.76×10^{-3}	4.52×10^{5}	4.13×10^{-2}

Table 5. A list of MAPE values of previous studies and this study.

Forecasting Models	MAPE (%)
Kang et al. [8]	3.84
Giudice et al. [9]	10.62
Plakandaras et al. [12]	2.151
Lee et al. [14]	4.54
Antipov and Pokryshevskaya [15]	13.95
Kusan et al. [18]	3.65
*LSSVR 1	1.676
**LSSVR 2	0.228

*LSSVR1: LSSVR models without attribute selection; **LSSVR 2: LSSVR models with attribute selection.

6. Conclusions

The decision to purchase real estate is undeniably very essential in the life of most adults. Thus, the appraisal and prediction of real estate can provide useful information to help facilitate real estate transactions. Real estate prices vary due to a wide variety of attributes. Machine learning models, including least squares support vector regression, classification and regression tree, general regression neural networks, and backpropagation neural networks, were used in this investigation to forecast real estate transaction prices with actual transaction data in Taichung, Taiwan. Genetic algorithms were employed to determine parameters of machine learning models. Empirical results revealed that attribute selection for machine learning models in this study does improve performances of four forecasting models in forecasting accuracy. With attribute selection, three machine learning models offer highly accurate predictions, and one machine learning model presents good prediction. Thus, four machines model with features selection used in this study are appropriate for forecasting real estate prices. Furthermore, the least squares support vector regression outperformed the other three forecasting models and obtains more accurate results than some previous studies in terms of MAPE. Thus, the least squares support vector regression with genetic algorithms is a feasible and promising machine learning technique in forecasting real estate prices.

For future work, diverse data types such as comments of real estate attributes, prices from social media, images from Google maps, and economic indicators are possible sources added as inputs for machine learning models to improve forecasting accuracy. In this study, general regression neural networks and backpropagation neural networks seems not to generate satisfying results when compared with results provided by the least squares support vector regression and the classification and regression tree. Thus, another potential opportunity for future research might be the use of deep learning techniques to forecast real estate prices.

Author Contributions: Conceptualization, P.-F.P.; data curation; W.-C.W.; formal analysis, P.-F.P. and W.-C.W.; funding acquisition, P.-F.P.; methodology, P.-F.P. and W.-C.W.; software, W.-C.W.; visualization, P.-F.P.; writing-original draft, review and editing, P.-F.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science and Technology, Taiwan under the Contract Number MOST 109-2410-H-260-023.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ahn, J.J.; Byun, H.W.; Oh, K.J.; Kim, T.Y. Using ridge regression with genetic algorithm to enhance real estate appraisal forecasting. *Expert Syst. Appl.* **2012**, *39*, 8369–8379. [CrossRef]
2. Gruma, B.; Govekar, D.K. Influence of Macroeconomic Factors on Prices of Real Estate in Various Cultural Environments: Case of Slovenia, Greece, France, Poland and Norway. *Procedia Econ. Financ.* **2016**, *39*, 597–604. [CrossRef]

3. Leamer, E.E. Housing is the Business Cycle. NBER Working Paper No. 13428. 2007. Available online: http://www.nber.org/papers/w13428 (accessed on 9 August 2020).
4. Beimer, W.; Maennig, W. Noise effects and real estate prices: A simultaneous analysis of different noise sources. *Transp. Res. Part D* **2017**, *54*, 282–286. [CrossRef]
5. Ferlan, N.; Bastic, M.; Psunder, I. Influential Factors on the Market Value of Residential Properties. *Inz. Ekon. Eng. Econ.* **2017**, *28*, 135–144. [CrossRef]
6. Singh, A.; Sharma, A.; Dubey, G. Big data analytics predicting real estate prices. *Int. J. Syst. Assur. Eng. Manag.* **2020**. [CrossRef]
7. Segnon, M.; Gupta, R.; Lesame, K.; Wohar, M.E. High-Frequency Volatility Forecasting of US Housing Markets. *J. Real Estate Finance Econ.* **2020**. [CrossRef]
8. Kang, H.; Lee, K.; Shin, D.H. Short-Term Forecast Model of Apartment Jeonse Prices Using Search Frequencies of News Article Keywords. *Ksce J. Civ. Eng.* **2019**, *23*, 4984–4991. [CrossRef]
9. Giudice, V.D.; Paola, P.D.; Forte, F. Using Genetic Algorithms for Real Estate Appraisals. *Buildings* **2017**, *7*, 31. [CrossRef]
10. Park, B.; Bae, P.J. Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data. *Expert Syst. Appl.* **2015**, *42*, 2928–2934. [CrossRef]
11. Bork, L.; Møller, S.V. Forecasting house prices in the 50 states using Dynamic Model Averaging and Dynamic Model Selection. *Int. J. Forecast.* **2015**, *31*, 63–78. [CrossRef]
12. Plakandaras, V.; Gupta, R.; Gogas, P.; Papadimitriou, T. Forecasting the U.S. real house price index. *Econ. Model.* **2015**, *45*, 259–267. [CrossRef]
13. Chen, Z.-H.; Tsai, C.-T.; Yuan, S.-M.; Chou, S.-H.; Chern, J. Big data: Open data and realty website analysis. In Proceedings of the 8th International Conference on Ubi-Media Computing, Colombo, Sri Lanka, 24–26 August 2015; pp. 84–88.
14. Lee, W.-T.; Chen, J.; Chen, K. Determination of Housing Price in Taipei City Using Fuzzy Adaptive Networks. In Proceedings of the International Multiconference of Engineers and Computer Scientists, Hong Kong, China, 13–15 March 2013.
15. Antipov, E.A.; Pokryshevskaya, E.B. Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics. *Expert Syst. Appl.* **2012**, *39*, 1772–1778. [CrossRef]
16. Kontrimas, V.; Verikas, A. The mass appraisal of the real estate by computational intelligence. *Appl. Soft. Comput.* **2011**, *11*, 443–448. [CrossRef]
17. Gupta, R.; Kabundi, A.; Miller, S.M. Forecasting the US real house price index: Structural and non-structural models with and without fundamentals. *Econ. Model.* **2011**, *28*, 2013–2021. [CrossRef]
18. Kusan, H.; Aytekin, O.; Özdemir, I. The use of fuzzy logic in predicting house selling price. *Expert Syst. Appl.* **2010**, *37*, 1808–1813. [CrossRef]
19. Cortes, C.; Vapnik, V. Support-vector networks. *MLear* **1995**, *20*, 273–297. [CrossRef]
20. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995.
21. Mukherjee, S.; Osuna, E.; Girosi, F. Nonlinear prediction of chaotic time series using support vector machines. In Proceedings of the IEEE Signal Processing Society Workshop, Amelia Island, FL, USA, 24–26 September 1997; pp. 511–520.
22. Müller, K.-R.; Smola, A.J.; Rätsch, G.; Schölkopf, B.; Kohlmorgen, J.; Vapnik, V. Predicting time series with support vector machines. In Proceedings of the International Conference on Artificial Neural Networks, Munich, Germany, 17–19 September 2019; pp. 999–1004.
23. Vapnik, V.; Golowich, S.E.; Smola, A.J. Support vector method for function approximation, regression estimation and signal processing. In Proceedings of the Advances Neural Information Processing System, Denver, CO, USA, 2–6 December 1997; pp. 281–287.
24. Suykens, J.A.; Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [CrossRef]
25. Fletcher, R. *Practical Methods of Optimization*; Wiley: Hoboken, NJ, USA, 1987; pp. 80–94.
26. Karush, W. Minima of Functions of Several Variables with Inequalities as Side Conditions. Master's Thesis, University of Chicago, Chicago, IL, USA, 1939.
27. Kuhn, H.W.; Tucker, A.W. Nonlinear programming. In Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probabilities, Berkeley, CA, USA, 31 July–12 August 1951; pp. 481–492.

28. Mercer, J. Functions of Positive and Negative Type and Their Connection with the Theory of Integral Equations. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **1909**, *209*, 415–446.
29. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. *Classification and Regression Trees*; Chapman and Hall, Wadsworth: New York, NY, USA, 1984.
30. Liu, Y.Y.; Yang, M.; Ramsay, M.; Li, X.S.; Coid, J.W. A Comparison of Logistic Regression, Classification and Regression Tree, and Neural Networks Models in Predicting Violent Re-Offending. *J. Quant. Criminol.* **2011**, *27*, 547–573. [CrossRef]
31. Parzen, E. On Estimation of a Probability Density Function and Mode. *Ann. Math. Statist.* **1962**, *33*, 1065–1076. [CrossRef]
32. Specht, D.F. A general regression neural network. *IEEE Trans. Neural Netw.* **1991**, *2*, 568–576. [CrossRef] [PubMed]
33. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [CrossRef]
34. Lin, C.T.; Lee, C.G. *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*; Prentice Hall: Upper Saddle River, NJ, USA, 1996.
35. Holland, J. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*; University of Michigan Press: Ann Arbor, MI, USA, 1975; pp. 439–444.
36. Lewis, C.D. *Industrial and Business Forecasting Methods*; Butterworth Scientific: London, UK, 1982.
37. Zhang, R.; Du, Q.; Geng, J.; Liu, B.; Huang, Y. An improved spatial error model for the mass appraisal of commercial real estate based on spatial analysis: Shenzhen as a case study. *Habitat Int.* **2015**, *46*, 196–205. [CrossRef]
38. Kato, T. Prediction in the lognormal regression model with spatial error dependence. *J. Hous. Econ.* **2012**, *21*, 66–76. [CrossRef]
39. Seya, H.; Yamagata, Y.; Tsutsumi, M. Automatic selection of a spatial weight matrix in spatial econometrics: Application to a spatial hedonic approach. *Reg. Sci. Urban. Econ.* **2013**, *43*, 429–444. [CrossRef]