

Research Article

Housing Value Forecasting Based on Machine Learning Methods

Jingyi Mu,¹ Fang Wu,² and Aihua Zhang³

¹ State Key Laboratory of Robotics and Systems (HIT), Harbin Institute of Technology, Heilongjiang 150001, China

² Research Institute of Intelligent Control and Systems, Harbin Institute of Technology, Heilongjiang 150001, China

³ College of Engineering, Bohai University, Jinzhou 121013, China

Correspondence should be addressed to Aihua Zhang; aihuazhangbhu@gmail.com

Received 28 June 2014; Accepted 15 July 2014; Published 4 August 2014

Academic Editor: Shen Yin

Copyright © 2014 Jingyi Mu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the era of big data, many urgent issues to tackle in all walks of life all can be solved via big data technique. Compared with the Internet, economy, industry, and aerospace fields, the application of big data in the area of architecture is relatively few. In this paper, on the basis of the actual data, the values of Boston suburb houses are forecast by several machine learning methods. According to the predictions, the government and developers can make decisions about whether developing the real estate on corresponding regions or not. In this paper, support vector machine (SVM), least squares support vector machine (LSSVM), and partial least squares (PLS) methods are used to forecast the home values. And these algorithms are compared according to the predicted results. Experiment shows that although the data set exists serious nonlinearity, the experiment result also show SVM and LSSVM methods are superior to PLS on dealing with the problem of nonlinearity. The global optimal solution can be found and best forecasting effect can be achieved by SVM because of solving a quadratic programming problem. In this paper, the different computation efficiencies of the algorithms are compared according to the computing times of relevant algorithms.

1. Introduction

Facing the upcoming era of big data, more and more people begin to engage in data analysis and mining. Machine learning [1], as a common means of data analysis, has gotten more and more attention. People of different industries are using machine learning algorithms to solve the problems based on their own industry data [2, 3]. Experts in the field of industry used machine learning in pattern recognition [4] and fault diagnosis [5, 6]. People in the field of economy began to use machine learning algorithms in economic modeling [7, 8]. The advantages of these algorithms were taken by the specialist in the field of aerospace in the aspect of classification and prediction [9]. Researchers in the field of construction combined machine learning methods with the professional domain knowledge of construction industry. Many intelligent systems are used in the construction industry, and lots of them have achieved good economic and social benefits. In general, the budget of construction project and benefit analysis of construction are usually gotten by the experience of the professionals and the construction of

traditional models. If the housing values can be accurately predicted, the government can make a reasonable urban planning. The historical housing price index was used by Malpezzi in 1999 to predict the changes of prices of 133 U.S. cities [10]. He thought that the price of the house was not randomly changed but followed certain rules. So, the prices can be partly predicted. Anglin predicted the real estate prices of Toronto by establishing a VAR model [11]. The results showed that the growth of the real estate prices is associated with unemployment and consumer prices. Some experts predicted the property's value by neural network.

Artificial neural network (ANN), which is constructed by a large number of neurons nodes and corresponding weights, is an artificial system to simulate the neural network of animals and plants in nature. Because of its good nonlinear characteristic, ANN can simulate the nonlinear function. But its accuracy is low and convergence speed is also not very ideal [12]. Since the ANN system is composed of a large number of individual neurons, the system has strong expansibility and evolutionary ability. Due to the existence of multiple equilibrium position, ANN may be trapped in local

minimum problem in optimization. It is difficult to obtain the global optimal value [13].

In order to find the global optimal solution, SVM method, which was founded on the statistical learning theory, was put forward in the 1990s [14, 15]. By solving a quadratic programming problem, SVM can find the global optimal solution. But when the samples are a lot, it leads to a higher complexity. For solving the problems of less number of samples, sample data is nonlinear and samples have high dimensions; SVM has great advantages. SVM not only has a strong learning ability but also has a strong generalization ability. SVM is mainly solving classification and regression problems. For the linear undivided sample data, SVM mapped the samples to a high-dimensional space to make them linear separable [16]. LSSVM is the improved results of SVM [17]. By changing the inequality constraints in the SVM into equality constraints, the original quadratic programming problem becomes a problem to solving system of linear equations. When compared with SVM, LSSVM not only reduced the complexity of the calculation but also improved the efficiency of calculation. In addition, the parameters that need to be adjusted in LSSVM are less. But LSSVM lost the sparse characteristic of SVM [18]. PLS is put forward by Wold to solve chemical sample analysis problems in the late 1960s. It has a lot of advantages in solving the problem of multiple variables, especially when all sample variables exhibit serious internal correlations. PLS algorithm is simple and has strong explanatory power. So it was gradually applied to other aspects besides the area of chemistry [19].

In this paper, the background and current situation of the application of machine learning methods are firstly introduced. The development of its application in construction and real estate value is also expounded. Then, several machine learning algorithms which are involved in this paper are introduced. The mathematics process of them is described in detail. It validates the models constructed by SVM, LSSVM and PLS with the real data of housing value of Boston suburb. According to the prediction results of these several methods, a discussion is made.

2. Algorithm

2.1. The Regression Algorithm of SVM. SVM was proposed in 1995 based on statistical learning theory [20]. Compared with the traditional machine learning methods, the machine learning algorithms at present stage are more rigorous in logic and more outstanding in generalization performance.

SVM is constructed on the basis of VC dimension theory and structure risk minimum principle, pursuing the best balance point between the learn ability and model complexity. In the problems of small sample, nonlinearity and high dimension, SVM has significant advantages. Small sample refers not to the absolute number of samples, but the number of samples which SVM algorithm requires is relatively small to the complexity of the problem. Solving the nonlinear problems is the core of SVM method. By the introduction of kernel function and slack variable SVM algorithm cleverly solved the problem of linear inseparable. Kernel function is a

function that satisfies the Mercer condition. The complexity of the calculation is effectively reduced by the kernel function. The reason why SVM has very big advantages in solving high dimension problems is that SVM does not need to use all of the samples in dealing with problems and only those support vectors are needed. The troubles in sample storage and computing are avoided [21].

SVM is mainly used to solve the problems of classification of the samples of different categories and the regression of the samples. The classification problem mainly refers to seeking a hyperplane in the higher dimensional space to separate out the samples of different categories. For SVM, the multiple classification can be solved via constructing two classifiers. SVM regression is to predict the concrete value on the basis of different sample characteristics [22, 23]. The training data set is defined as $(x_i, y_i), i = 1 \cdots m$, where x_i is the input data and y_i is the corresponding output data.

The regression function can be expressed as follows:

$$y = (z^* \cdot x_i) + l, \quad (1)$$

where z^* represents weight vector and l is a normal. Both of these parameters can be gotten from the function as follows:

$$\begin{aligned} \min_{z^*, b, \xi^{(*)}} \quad & \frac{1}{2} \|z^*\|^2 + F \sum_{i=1}^q (s_i + s_i^*) \\ \text{s.t.} \quad & \begin{cases} ((z^* \cdot x_i) + l) - y_i \leq \varepsilon + s_i, & i = 1, \dots, q \\ y_i - ((z^* \cdot x_i) + l) \leq \varepsilon + s_i^*, & i = 1, \dots, q \\ s_i^* \geq 0, & i = 1, \dots, q, \end{cases} \end{aligned} \quad (2)$$

where ε is the insensitive loss function, F represents the punishment coefficient, and s_i^* is the slack variable.

The introduction of kernel function can solve the problem that the nonlinear mapping function cannot be expressed. According to the kernel function, the inner product of higher dimension feature space can be obtained. The nonlinear regression function can be expressed as

$$\begin{aligned} y &= \sum_{i=1}^q (\tilde{\alpha}'_i - \tilde{\alpha}_i) K(x_i, x) + \bar{l}, \\ \text{s.t.} \quad & \begin{cases} \sum_{i=1}^q (\tilde{\alpha}'_i - \tilde{\alpha}_i) = 0 \\ 0 \leq \tilde{\alpha}'_i, \quad \tilde{\alpha}_i \leq P, \quad i = 1, \dots, q, \end{cases} \end{aligned} \quad (3)$$

where $\tilde{\alpha}^{(l)} = (\tilde{\alpha}_1, \tilde{\alpha}'_1, \dots, \tilde{\alpha}_q, \tilde{\alpha}'_q)^T$ is the solution.

Like other multivariate statistical models, the selection of each variable's value in the model has deep influence to their model's performance, such as the type of kernel function and the corresponding kernel parameters. In SVM regression model, regularization parameter and ε -insensitive loss function played a crucial role. Regularization parameter determines the balance between maximum classification interval and minimizing training error. If regularization parameter is chosen inappropriately, the model may appear overfitting and underfitting. ε is also an important parameter, in which value depends on the type of noise.

2.2. *The Regression Algorithm of LSSVM.* LSSVM is put forward by Suykens et al. [24]. The purpose of it is to solve some problems that exist in SVM algorithm, such as the selection of hyperplane parameters, and the matrix scale is highly affected by the number of training samples when solving a quadratic programming problem. Suykens starts from the loss function, using the two-norm in the objective function and replacing inequality constraints in SVM algorithm by equality constraints. Using Kuhn-Tucker conditions can get the solution of this set of equations. When the data size reaches a certain extent, the scale caused by solving a quadratic programming problem is very large. Some traditional methods are difficult to apply in such a large data size. By solving the linear equations, LSSVM not only reduces the difficulty of solution but also promotes the speed to solve the problem. So it is more suitable for applying to solve the problems of data on a large scale [25].

Although LSSVM may not necessarily be able to obtain the global optimal solution, it still can get a high recognition rate. The square of error is used in objective function of LSSVM optimization problem, that is the Lagrange Multipliers are proportional to the error term. The direct consequence is that the ultimate decision function has relationship with all of the samples, which represents that LSSVM loses the sparse characteristic of SVM solution [26].

The training samples are described as $(x_j, y_j), j = 1 \cdots m$. x_j is the input data. Each column represents a feature. y_j is the output data. Owing to these training data, the regression model can be described as follows:

$$y = z^T \psi(x) + l, \tag{4}$$

where $\psi(x)$ represents the nonlinear mapping function, z represents the weight factor, and l is the invariable coefficient. The cost function can be described as follows:

$$\min J(z, e) = \frac{1}{2} z^T z + r \frac{1}{2} \sum_{j=1}^m e_j^2 \tag{5}$$

$$\text{s.t. } y_j = z^T \psi(x_j) + l + e_j,$$

where e_j represents random error and r represents regularization parameter. Lagrange function is introduced to solve this equation

$$L(z, l, e, \tilde{\alpha}) = J(z, e) - \sum_{j=1}^m \tilde{\alpha}_j \{z^T \psi(x_j) + l + e_j - y_j\}, \tag{6}$$

where $\tilde{\alpha}_j$ represents Lagrange multiplier. The solution of (6) can be obtained by partially differentiating

$$\frac{\partial L}{\partial z} = 0 \longrightarrow z = \sum_{j=1}^m \tilde{\alpha}_j \psi(x_j),$$

$$\frac{\partial L}{\partial l} = 0 \longrightarrow \sum_{j=1}^m \tilde{\alpha}_j = 0,$$

$$\frac{\partial L}{\partial e_j} = 0 \longrightarrow \tilde{\alpha}_j = r e_j,$$

$$\frac{\partial L}{\partial \tilde{\alpha}_j} = 0 \longrightarrow z^T \psi(x_j) + l + e_j - y_j = 0. \tag{7}$$

Equation (7) can be described as

$$\begin{pmatrix} 0 & \tilde{1}^T \\ \tilde{1} & \Omega + r^{-1}I \end{pmatrix} \begin{pmatrix} l \\ \tilde{\alpha} \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}. \tag{8}$$

The solution of l and $\tilde{\alpha}$ can be obtained from (8). After introducing the kernel function, the objective function could be rewritten as

$$y(x) = \sum_{j=1}^m \tilde{\alpha}_j K(x, x_j) + l. \tag{9}$$

2.3. *Partial Least Squares Regression Algorithm.* In 1983, PLS regression was put forward for the first time. As a new method of multivariate statistical data analysis, it has many advantages that traditional multivariate statistical analysis methods do not have [27].

PLS can find the best function matching with the original data accordingly to minimize the sum of the squares of error. Although the independent variables have multiple correlation, PLS is still able to build the model [28]. All of the independent variables will be contained in the final model of PLS regression. And maximum information will be extracted from the original data, which ensures the accuracy of the model. PLS model is easier to distinguish the noise from normal information. When the number of variables is greater than the number of sample points, PLS algorithm is often chosen to construct model. PLS can construct regression model, simplify the data structure, and analyze the correlation at the same time [29].

The mathematical derivation processes of PLS regression are to be described as follows.

The input matrix should be normalized at first. Normalization can make the follow-up data processing more convenient and prevent the predicting accuracy to be influenced by the existence of singular samples. Consider $X = [x_1, x_2, \dots, x_b]_{a \times b}$ and $Y = [y]_{a \times 1}$,

$$x'_{mn} = \frac{x_{mn} - \bar{x}}{d_x}, \quad m = 1, \dots, a; \quad n = 1, 2, \dots, b, \tag{10}$$

$$M_0 = (x'_{mn})_{a \times p}, \quad N_0 = \left(\frac{y_m - \bar{y}}{d_y} \right)_{a \times 1},$$

where \bar{x} represents mean value of X_n , d_x represents standard deviation of X_n , \bar{y} represents mean value of Y , and d_y represents standard deviation of Y .

h_1 is the main ingredient, which is extracted from M_0 ; $h_1 = M_0 w_1$. The more h_m includes the information of M_0 ,

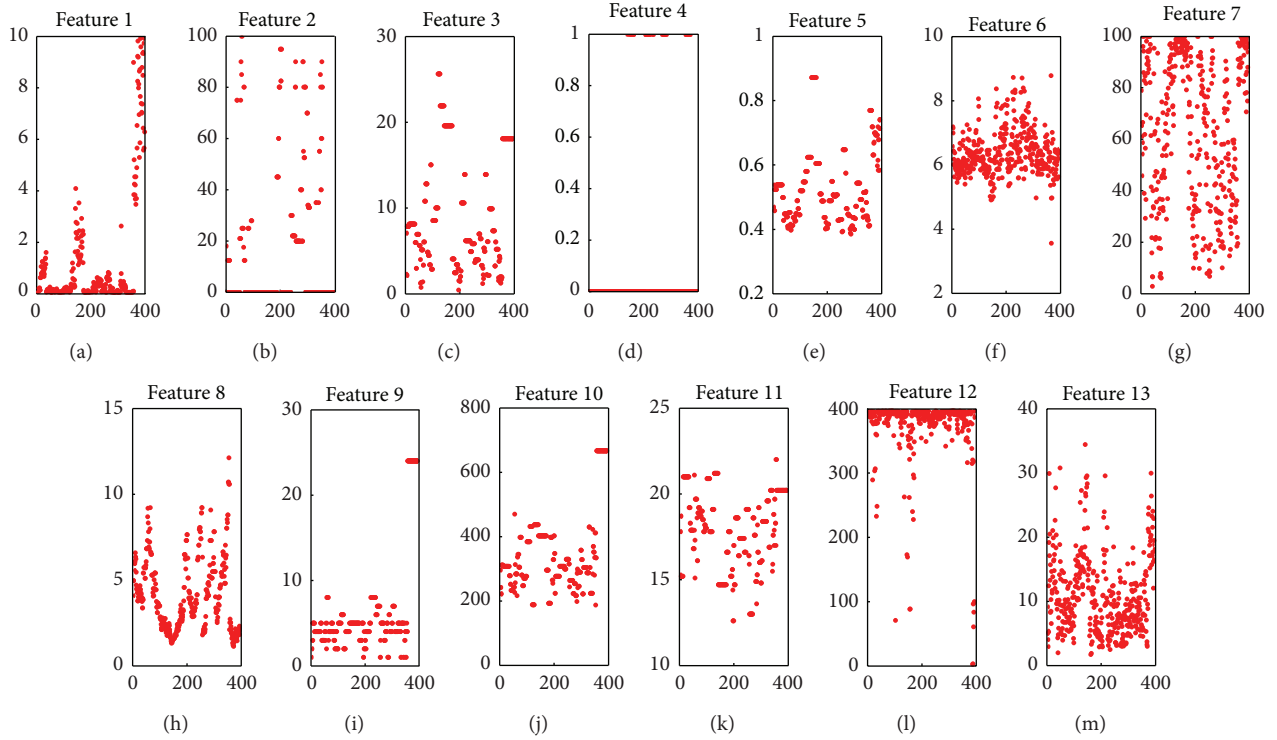


FIGURE 1: Features of the training data.

the better h_m can be on behalf of M_0 . After this, h_1 has the major information of y ,

$$\begin{aligned} w_1 &= \frac{M_0' N_0}{\|M_0' N_0\|}, \quad \|w_1\| = 1, \\ M_0 &= h_1 l_1' + M_1, \\ N_0 &= h_1 p_1 + N_1, \end{aligned} \quad (11)$$

where l_1 and p_1 are two parameters,

$$l_1 = \frac{M_0' h_1}{\|h_1\|^2}, \quad p_1 = \frac{N_0' h_1}{\|h_1\|^2}. \quad (12)$$

The residual matrix can be expressed as follows:

$$\begin{aligned} M_1 &= M_0 - h_1 l_1', \\ N_1 &= N_0 - h_1 p_1. \end{aligned} \quad (13)$$

According to the above steps, if the regression equation has achieved sufficient accuracy, algorithm will continue to perform the following steps. But when the precision does not meet the requirement, more ingredients should be picked up from the residual matrix.

Sufficient ingredient can be extracted after i rounds; the equation has achieved the precision, $i = 2, \dots, k$, where k represents the number of main ingredients. The regression equation of N_0 is described as follows:

$$\tilde{N}_0 = p_1 h_1 + p_2 h_2 + \dots + p_k h_k. \quad (14)$$

Since M_0 can be expressed by the combination of h_1, \dots, h_k , \tilde{N}_0 is expressed as follows:

$$\tilde{N}_0 = p_1 M_0 w_1' + \dots + p_k M_k w_k', \quad (15)$$

where w_i' is $w_i' = \prod_{n=1}^{i-1} (E - w_n l_n')$ and E is the unit matrix.

The regression equation of Y on X can be obtained by an inverse process of standardization.

3. Experimental

In this paper, the data is selected from UCI data sets, which can be downloaded from the Internet. Housing value of Boston suburb can be measured through the data of 13 features. These features include per capita crime rate by town, proportion of nonretail business acres per town, and index of accessibility to radial highways.

Housing value of Boston suburb is analyzed and forecast by SVM, LSSVM, and PLS methods and the corresponding characteristics. After getting rid of the missing samples from original data set, 400 samples are treated as training data and 52 samples are treated as test data. Housing value of the training data can be seen in Figure 2. All features of the training data can be seen in Figure 1. The predicted results are shown as follows. The ordinate means the median value of owner-occupied homes in the \$1000's and the abscissa represents different sample points. The blue line represents the real data values and the red line represents predicted values through different machine learning models. According to the comparative results between the real values

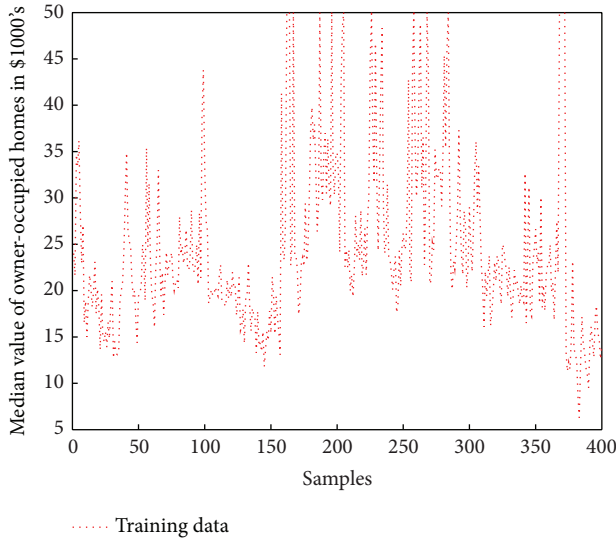


FIGURE 2: Housing value of the training data.

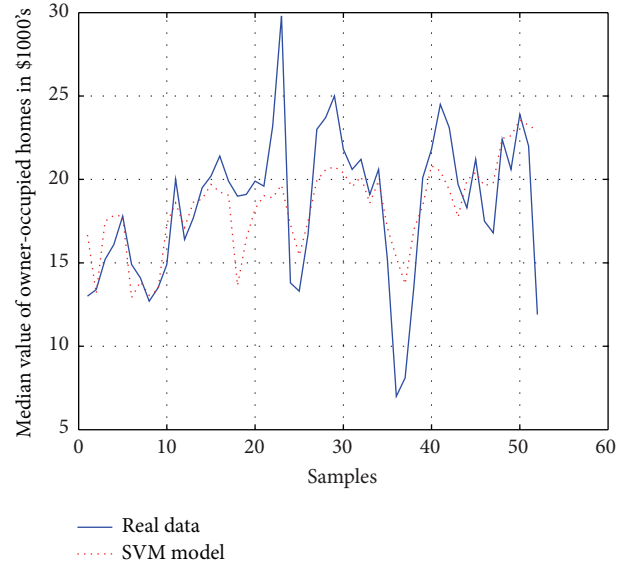


FIGURE 3: Forecasting results by SVM.

and predicted values, the effects of different methods can be compared.

After constructing SVM model by the training samples, housing value can be forecast on the basis of testing samples. From the contrast between real data and forecasting data in Figure 3, it can be seen that although forecasting data have certain deviations from the real data, they can also reflect the change trend of different samples anyway. The mean square error of estimating data, which is obtained by SVM, is 10.7373. The running time of SVM algorithm is 0.4610 s.

LSSVM algorithm is used to predict the housing value of Boston suburb. As shown in Figure 4, several forecasting data have relative great deflection to the real data. The mean square error of forecast data is 15.1310, which shows that the accuracy of SVM is bigger than that of LSSVM. The running time of LSSVM algorithm is 20.3730 s. Since a parameters optimization step is joined in LSSVM program, the calculation time of overall program is longer than that of SVM method. After removing the parameter optimization process, the operation time is 0.3460 s, which suggests that LSSVM has a higher computational efficiency. But the corresponding prediction effect will be worse.

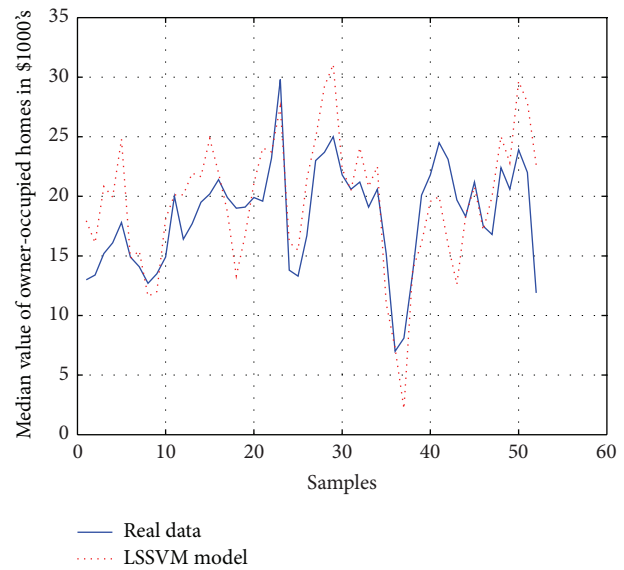


FIGURE 4: Forecasting results by LSSVM.

As shown in Figure 5, PLS is used for predicting the homes' value of Boston suburb. The predicting situation is not very ideal. There is a big deviation with the real value. The mean square error of predictive value is 25.0540. The algorithm's running time is 0.7460 s. For this nonlinear data sets, the forecasting ability of PLS is obviously worse than that of SVM and LSSVM.

According to the predicting results of home's value of Boston suburb, SVM has a higher prediction accuracy than LSSVM and PLS. Because of LSSVM simplified mathematical mechanism of SVM, its computation efficiency is the highest. Due to the presence of strong nonlinearity about the home's value of Boston suburb in the data set, the forecast result of PLS algorithm is not very ideal and the computation efficiency is very low.

4. Conclusion

In this paper, SVM, LSSVM, and PLS algorithms are used in the field of construction to predict the housing value. According to multiple characteristics, the housing value of Boston suburb is forecasted. The models of several machine learning methods should be constructed and analyzed at first and then combined with the corresponding characteristics of testing data to predict the housing value. The prediction results of various machine learning approaches are not the same. Aiming at the nonlinear data, SVM and LSSVM have better prediction effect, learning ability, and generalization ability than PLS. The prediction effect of SVM is superior to that of LSSVM. LSSVM is remoulded on the basis of SVM

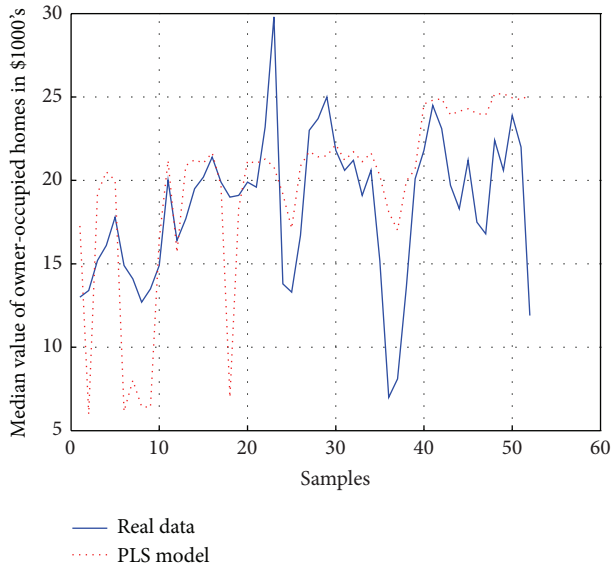


FIGURE 5: Forecasting results by PLS.

mathematical process. The quadratic programming problem in SVM is transformed into solving an equation system in LSSVM, which leads to the fact that the computation complexity is reduced and the calculation efficiency of LSSVM algorithm is higher than SVM algorithm. Compared with PLS, SVM and LSSVM are more suitable for the nonlinear field. Due to the simplicity of algorithm, PLS algorithm is more suitable for the linear system. At this stage, PLS is widely used in industrial and other fields.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This present work was supported partially by the Open Foundation of State Key Laboratory of Robotics and Systems (HIT). The authors highly appreciate the above financial supports.

References

- [1] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- [2] S. Yin, G. Wang, and H. Karimi, "Data-driven design of robust fault detection system for wind turbines," *Mechatronics*, vol. 24, no. 4, pp. 298–306, 2014.
- [3] S. Yin, X. Yang, and H. R. Karimi, "Data-driven adaptive observer for fault diagnosis," *Mathematical Problems in Engineering*, vol. 2012, Article ID 832836, 21 pages, 2012.
- [4] P. Bhagat, *Pattern Recognition in Industry*, Elsevier, New York, NY, USA, 2005.
- [5] A. Widodo and B. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2560–2574, 2007.
- [6] S. Yin, X. Li, H. Gao, and O. Kaynak, "Data-based techniques focused on modern industry: an overview," *IEEE Transactions on Industrial Electronics*, 2014.
- [7] C. L. Huang, M. C. Chen, and C. J. Wang, "Credit scoring with a data mining approach based on support vector machines," *Expert Systems with Applications*, vol. 33, no. 4, pp. 847–856, 2007.
- [8] S. Ding, S. Yin, K. Peng, H. Hao, and B. Shen, "A novel scheme for key performance indicator prediction and diagnosis with application to an industrial hot strip mill," *IEEE Transaction on Industrial Informatics*, vol. 9, no. 4, pp. 2239–2247, 2013.
- [9] W. J. Staszewski, S. Mahzan, and R. Traynor, "Health monitoring of aerospace composite structures—active and passive approach," *Composites Science and Technology*, vol. 69, no. 11–12, pp. 1678–1685, 2009.
- [10] S. Malpezzi, "A simple error correction model of housing price," *Journal of Housing Economics*, vol. 8, no. 1, pp. 27–62, 1999.
- [11] P. Anglin, "Local dynamics and contagion in real estate markets," in *Proceedings of the International Conference on Real Estates and Macro Economy*, pp. 1–31, Beijing, China, 2006.
- [12] A. Abraham, *Artificial Neural Networks*, Handbook of Measuring System Design, 2005.
- [13] B. Samanta and K. R. Al-Balushi, "Artificial neural network based fault diagnostics of rolling element bearings using time-domain features," *Mechanical Systems and Signal Processing*, vol. 17, no. 2, pp. 317–328, 2003.
- [14] H. Yoon, S. Jun, Y. Hyun, G. Bae, and K. Lee, "A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer," *Journal of Hydrology*, vol. 396, no. 1–2, pp. 128–138, 2011.
- [15] S. Yin, S. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Transactions on Industrial Electronics*, 2014.
- [16] S. Shalev-Shwartz, Y. Singer, and N. Srebro, "Pegasos: primal estimated sub-gradient solver for svm," *Mathematical Programming*, vol. 127, no. 1, pp. 3–30, 2011.
- [17] L. Yu, H. Chen, S. Wang, and K. K. Lai, "Evolving least squares support vector machines for stock market trend mining," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 1, pp. 87–102, 2009.
- [18] T. van Gestel, J. de Brabanter, and B. de Moor, *Least Squares Support Vector Machines*, vol. 4, World Scientific, Singapore, 2002.
- [19] S. Yin, G. Wang, and X. Yang, "Robust PLS approach for KPI related prediction and diagnosis against outliers and missing data," *International Journal of Systems Science*, vol. 45, no. 7, pp. 1375–1382, 2014.
- [20] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [21] C. Chang and C. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, 2011.
- [22] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, "Support vector machines for wind speed prediction," *Renewable Energy*, vol. 29, no. 6, pp. 939–947, 2004.
- [23] A. S. Naik, S. Yin, S. X. Ding, and P. Zhang, "Recursive identification algorithms to design fault detection systems," *Journal of Process Control*, vol. 20, no. 8, pp. 957–965, 2010.
- [24] J. A. K. Suykens, T. van Gestel, J. de Brabanter, B. de Moor, and J. Vandewalle, *Least Squares Support Vector Machines*, World Scientific Publishing, Singapore, 2002.

- [25] D. Wu, Y. He, S. Feng, and D. Sun, "Study on infrared spectroscopy technique for fast measurement of protein content in milk powder based on LS-SVM," *Journal of Food Engineering*, vol. 84, no. 1, pp. 124–131, 2008.
- [26] A. Borin, M. F. Ferrão, C. Mello, D. A. Maretto, and R. J. Poppi, "Least-squares support vector machines and near infrared spectroscopy for quantification of common adulterants in powdered milk," *Analytica Chimica Acta*, vol. 579, no. 1, pp. 25–32, 2006.
- [27] J. Henseler, C. M. Ringle, and R. R. Sinkovics, "The use of partial least squares path modeling in international marketing," *Advances in International Marketing*, vol. 20, no. 1, pp. 277–319, 2009.
- [28] P. P. Roy and K. Roy, "On some aspects of variable selection for partial least squares regression models," *QSAR and Combinatorial Science*, vol. 27, no. 3, pp. 302–313, 2008.
- [29] O. Gotz, K. Liehr-Gobbers, and M. Krafft, "Evaluation of structural equation models using the partial least squares (PLS) approach," in *Handbook of Partial Least Squares*, pp. 691–711, Springer, Berlin, Germany, 2010.