MATEMATIKA: MJIAM, 2021, Volume 37, Number 3, 135–146

A Forecasting Cycling Distance Based on Personal Health Data with Hybrid GA-SVR Approach

[1]Pawnwipa Meeklueb, [2]Busayamas Pimpunchat* and [3]Benjawan Rodjanadid
[1,2] Department of Mathematics, School of Science
King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
[3]School of Mathematics, Institute of Science
Suranaree University of Technology, Nakhon Ratchasima, 30000, Thailand
*Corresponding author: busayamas.pi@kmitl.ac.th

Article history
Received: 21 August 2021
Received in revised form: 11 November 2021
Accepted: 21 November 2021
Published online: 15 December 2021

Abstract The aim of this study is to forecast the optimal cycling distance of 30 minutes approximately for physically constrained individuals, especially those having a congenital disease or elderly person. Data illustrations are collected from having 94 people who cycled in Ladkrabang by questionnaires. The research focuses on forecasting optimal distances by exploiting Support Vector Regression based on Genetic Algorithm (GA-SVR) that will be employed to improve the accuracy of the numerical study. We offer performance comparisons of our model against Multiple Linear Regression (MLR) and Support Vector Regression (SVR) algorithms. The experimental results demonstrate that GA-SVR outperforms MLR and SVR based on the Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). This GA-SVR model is proven to be an effective approach to predict cycling distances.

Keywords Optimal Cycling Distance; Health; Support Vector Regression; Genetic Algorithm; Multiple Linear Regression.

Mathematics Subject Classification 6211; 62P05; 62J99.

1 Introduction

Human beings are becoming increasingly interested in exercising nowadays since there are various forms of exercise, either with or without equipment. The merit of exercising is keeping the body healthy, but overexertion can lead to injuries. Settling with knowing exercise limits is recommended, and also it can contribute to reducing risks of adverse effects later on. Cycling is a popular exercise for people of all ages, yet some groups have limitations with physical activities, i.e. somebody with high blood pressure, an elderly person with diabetes, or groups at risk of developing coronary artery disease. Adequate designs of exercises combined with a healthy diet can reduce the risk of developing the disease [1, 2]. Physical activity is prompted for adults aged 18-64 years by performing moderate-intensity aerobic exercise for at least 150 minutes per week, or sets of 30 minutes, five times per week [3]. Much of those with physical limitations can only be flexible 2-3 times a week [4]. According to research and markets [5], the world's leading source for international market research reports and market data, it is stated that fitness equipment sales have gone up to 170%. Since gyms are required to close, it prompts many people to buy exercise equipment, i.e. exercise bikes, treadmills, and rowing machines. In a stationary bicycle, there is usually a screen displaying stats, i.e. the person's heart rate, distance while cycling, etc. It will be better if we can forecast the appropriate distance for each person. This is consistent with the current situation due to the widespread disease of the coronavirus outbreak, making people unable to exercise in a public space, where as a result, exercising at home can be an alternative.

For earlier studies, exercise can have beneficial effect on the elderly people [6] and cycling reduces the risk of various diseases in both children and adults [7]. The most common forecasting techniques are the Multiple Linear Regression (MLR) and the Support Vector Regression (SVR) methods, which are regression analyses between vectored inputs and output variables using machine learning principles [8]. These methods provide high accuracy [9-12], but the SVR has limitations in its functionality since the user has to configure the parameters accordingly to get an ideally powerful SVR. Therefore, people prefer to use the SVR with other methods to get more efficiency, Yang *et al.* [13], have developed a model that predicts consumers' affective responses (CARs) for producing a form design. The predictive performance of the SVR with Real-Coded Genetic Algorithm (SVR–RCGA) is compared to that of SVR with 5-fold cross-validation (SVR–5FCV) and a back-propagation neural network (BPNN) with 5-fold cross-validation (BPNN–5FCV). The experimental results using the data sets on mobile phones and electronic scooters show that SVR performs better than BPNN. MLR methods have been proposed by collecting data from cyclists on a cycling track as a sample for forecasting the distance of a stationary bike.

To find the appropriate parameters in the SVR, we propose the Genetic Algorithm based on Support Vector Regression (GA-SVR). GA-SVR can be applied to solve forecasting a wide variety of problems and even enhance it, i.e. Zangooei *et al.* [14], displayed the diagnosis of disease by the prognosis of four diseases, namely breast cancer, liver disease, hepatitis, and diabetes by utilizing SVR and Non-dominaetd Sorting Genetic Algorithm (NSGA-II). The results of the accuracy compared to other methods have found that the application of the SVR method with NSGA-II has higher accuracy than any other method. Tang *et al.* [15], proposed an unprecedented hybrid model based on a fuzzy information granulation that is integrating models of the GA-SVR and Autoregressive Integrated Moving Average (ARIMA) in the predicting of the Producer Price Index (PPI). Their research showed that the PPI value predicted by this hybrid model is more accurate than that predicted by other models. Guo *et al.* [16], forecasted the loud traffic of Xinjiang in holidays. All of the previous studies can be referred to [17, 18], and other related references. Numerical comparisons of this paper between MLR, SVR, and GA-SVR with measurements of Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE) are discussed in section 3 of this paper. In this paper, we are keen on the optimal cycling distance in 30 minutes so that users might utilize it as a guide to avoid excessive exercise effects that are possibly causing more peril than nifty.

2 Materials and Methods

The variables mentioned above are used for modeling both SVR and GA-SVR methods. There are 94 data in this study, divided into two parts, training data, and validation data. The numbers of training and validation data are divided into 80:20 ratio respectively. This split data is used in MLR, SVR, and GA-SVR methods to compare a model performance.

2.1 Previous Study

In 2019, Yotharak *at el.* [19], created the model using the MLR method with the following variables: Cycling distance in 30 minutes, gender, age, weight, height, cycling frequency, and heart rate. All of these variables are found to have correlations due to the MLR method. The data can be divided as follows: male aged 20-40 years, female aged 20-40 years, male aged 41-60 years, female aged 41-60 years, male aged 61-80 years and female aged 61-80 years and are then forecasted to get a total of 6 equations. This is because when using the data before sharing to forecast, the accuracy is increasingly low, but after the data is divided, it can be forecasted with better accuracy but is still not decent. Our past study used MLR attempts to model the relationship between the more explanatory variables and a response variable by fitting a linear equation to the observed data. Every value of the independent variable x is associated with a value of the dependent variable y. The multiple linear regression model is shown below

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + +a_5x_5 + b, \tag{1}$$

where each x_1, x_2, x_3, x_4 and x_5 represents ages, weight, height, cycling frequency and heart rate while y is the cycling distance for 30 minutes. The constants to be determined are $a_1, \ldots, a_5$ which are the coefficients of the model, a_0 as the intercept and b is the forecast error.

2.2 Data

This study examines the analysis of the optimal distance for cycling. The instrument used for this data collection is questionnaires, which is a collection of data from people who cycled at the "Happy and Healthy Bike Lane" in Ladkrabang, Bangkok, Thailand, with a total of 94 samples, aged between 20-80 years. Table 1 represents examples of the data collected including Gender (let the number 1 reflects the man, 2 with the woman), Age in year, Weight in kilogram, Height in centimeter, Dis30min is the distance for 30 minutes (kilometers), Freq is cycling frequency (times per week) and HR is heart rate (beats per minute) which is recorded at the moment of exhaustion.

2.3 Support Vector Regression

We recall tools of the Support Vector Machine (SVM) that is useful for our main discussions. For developing a SVR model, the SVM should be evaluated. In SVM, the largest margin between one class and another one is used as the basis for deciding the optimal hyperplane, and the data vectors at the edge of each class are called support vectors [20], hyperplanes are split into 2 types. Figure 1(a) shows a hyperplane capable of dividing two classes of data correctly, the hard-margin method is used. Figure 1(b) shows a hyperplane that cannot properly divide the two classes of data and use of the soft-margin method [21].

Table 1: Sample information of people riding bicycles

Gender	Age	Weight	Height	Dis30min	Freq	HR
1	27	88	181	12	3	115.8
1	24	76	173	25	3	117.6
1	35	76	170	10	2	111
1	26	68	180	15	3	170
1	25	67	169	15	1	168

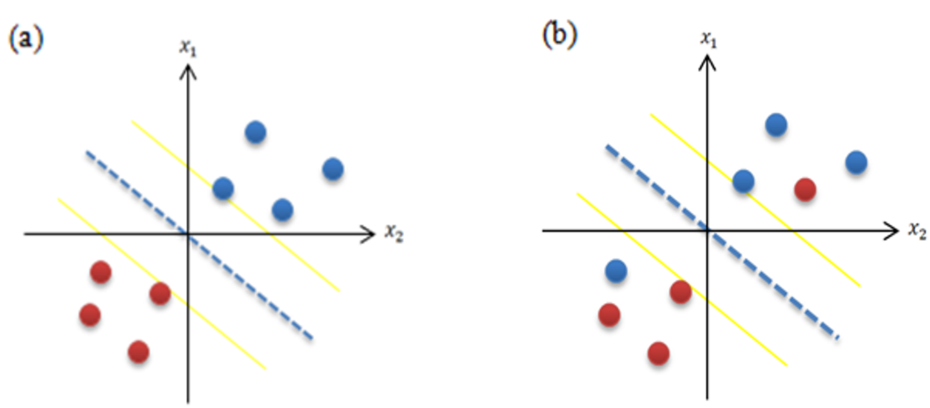

Figure 1: (a) hard-margin SVM and (b) soft-margin SVM

SVM is proposed by Vapnik [22], who thought that theoretically, a linear function f exists to define the nonlinear relationship between the input and output data in a high-dimensional-feature space. Such a method can be used to solve the function concerning fitting problems [23]. The SVR uses the same principle as SVM for classification, but regression problems. The main idea is always the same: to minimize error, individualizing the hyperplane which maximizes the margin, keeping in mind that part of the error is tolerated [24]. In the case of finding a linear relationship between the input vector X of n dimensions ($X \in R^n$) and the output variables ($y \in R$) of the training data set [8], the SVR regression equation is similar to SVM hyperplane equations because the SVR equation is adapted from SVM. The SVR model is displayed as following

$$f(x) = w^T x + b, \tag{2}$$

where w and b denoted as a weight vector and a bias, respectively. This leads to an optimization problem that minimizes the objective function as

$$\text{Minimize } \frac{1}{2}\left\|w^2\right\| \text{ subject to } y_i - (w^t x_i) - b \leq \varepsilon,\ (w^t x_i) + b - y_i \leq \varepsilon. \tag{3}$$

It is implicitly assumed that there is a function f that approximates all the pairs (x_i, y_i) with ε precision. In the case where this is not possible, or a specified error can be tolerated, slack variables ξ_i, ξ_i^* can be stated according to the formulation [25] as following

$$\text{Minimize}\, \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*)\, \text{subject to}\, y_i - (w^t x_i) - b \le \varepsilon + \xi_i,\ (w^t x_i) + b - y_i \le \varepsilon + \xi_i^*, \tag{4}$$

where C is a constant which is a parameter for determining the error quantity. In some cases, we cannot directly find the value w. The equation (4) above can solve the dual formulation by applying the Lagrange Multipliers with the Karush-Kuhn-Tucker (KKT) optimality condition. After solving the equation for the value of w, then nonlinear hyperplane regression functions can be formulated in equation (5) as follows

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(x_i, x) + b, \tag{5}$$

where α_i and α_i^* are Lagrange's multipliers and $K(x_i, x)$ is a kernel function [25], employed in cases of nonlinear regressions, can be exploited and relegated input vectors to higher dimensions feature space. The kernel function commonly used in SVR formulation are as follows:

Linear: $$K(x_i, x_j) = x_i^T x_j, \tag{6}$$

Polynomial: $$K(x_i, x_j) = \left(\gamma x_i^T x_j + r\right)^d, \tag{7}$$

Radial Basis Function (RBF): $$K(x_i, x_j) = e^{-\gamma\|x_i - x_j\|^2},\ \gamma > 0, \tag{8}$$

Sigmoid: $$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r). \tag{9}$$

For kernel functions, the parameters γ, r and d in equations (7) until (9), are constants depending on their suitability. In forecasting process, the input data, i.e., a distance for 30 minutes and related factors, for instance, gender, age, weight, height, cycling frequency, and heart rate, severally. The forecasting part, SVR model utilized the RBF kernel function from the previous research [26], which suggests the RBF is usually superior to other kernel. The calculation procedure will call the function in R programming. The forecasting result has gotten a cycling distance in 30 minutes. This research evaluates the forecast error values with RMSE and MAPE. We offered a GA-SVR method, (C, ε, γ) as crucial parameters, in order to improve values to achieve a better and highest accuracy than SVR method.

2.4 Genetic Algorithm Based on Support Vector Regression

To optimize model based on GA-SVR, a GA is needed to optimize the parameters (C, ε, γ) in the SVR model where C as a penalty coefficient, γ for a kernel function coefficient, and ε is the insensitive coefficient, which has the following main steps [27]:

1. Constructing the chromosome assemblage (C, ε, γ) and formulating the fitness calculation function of the genetic algorithm;
2. Determining the parameters of selection, crossover, mutation, and so on in the GA, and setting the iterative termination condition of the algorithm;
3. Initializing the GA and generating the initialization population;
4. Calculating individual fitness in chromosome populations;

5. Generating next-generation chromosomes through selection, crossover, variation, etc.

Exercise for 30 minutes is a suitable time for exercise. The focus now is on using time as one of the limitations in cycling distance forecasting, as a guide for those interested in knowing the amount of cycling distance for 30 minutes (kilometers). In addition, for those with underlying medical conditions, this model can be used to reduce the risk of injury from excessive exercise. The modeling process for comparing cycling distance forecasting with MLR, traditional SVR and GA-SVR models, as shown in Figure 2, is summarized as follows.

1. (Input Data). Call data sets consisting of gender, age, weight, height, cycling frequency per week, heart rate, and distance;
2. Separating the data in step 1 into trained data (80%) and validated data (20%);
3. (MLR and SVR). Applying for the same data sets to build the cycling model in optimal distance with MLR and traditional SVR methods. Finally, we can find the forecast data of two methods;
4. (GA-SVR). Three parameters are (C, ε, γ), for random initial values case, then apply for it for updating the SVR model;
5. (GA-SVR model). As for the GA-SVR model, random initial values of the three parameters (C, ε, γ) then apply them to update the SVR model calculating the fitness function by employing RMSE:
   a. The stop condition for the calculation is the maximum number of rounds (10, 50, 100, and 500 rounds of repeated sequences). If they do not yet reach the maximum rounds yet, continue with the calculation inside the loop and through the selection, crossover, and mutation processes;
   b. The results are parameter sets of a new population. These parameters can be put through an update again to see if they reach the maximum rounds in the set condition;
   c. If so, bring them out of the loop to get the most optimal parameters. Then, put the obtained parameters by utilizing a GA into the SVR model. The obtained result will be the GA-SVR cycling distance model;
   d. To find data forecasts of the GA-SVR method;
6. (Measure Performance). To compare the effectiveness of the three models with RMSE and MAPE.

3 Results and Discussions

This part is the main section of creating an effective machine learning model that uses the OS Intel Core i5-5200U, 4 GB RAM, 64 bit for numerical simulations by R programming. To evaluate the performance of the forecasting models, RMSE and MAPE were occupied to evaluate testing performance, which is defined by RMSE equation as

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}. \tag{10}$$

and a MAPE as follows

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right|. \tag{11}$$

From equations (10) and (11), y_t and $\hat{y}_t$ named actual and forecasting values, while n indicated the length of the validation subset. This research compared the performance of the 3 models with RMSE and MAPE error measurements from the validation data, MLR, and SVR as shown in Table 2.

Table 2: Comparisons of the efficiency of MLR and SVR

Method	Time (minutes)	RMSE	MAPE
MLR	0.010	1.630	0.120
SVR	0.001	1.622	0.110

We found that the SVR method had given more accuracy than MLR. The results are decent for acceptance, but the group study was compulsory to be examined with limited health conditions, so the efficiency needs to be improved. As a result, the GA method was used to combine with the SVR to achieve the closest forecast, the comparison results are shown in Figure 3. The figure shows the performance comparison of the 3 methods with the GA-SVR method is close to original values. In MLR, the correlation of variables was examined before the forecasting process, then the variables relating to each other were taken in forecasting both traditional SVR and GA-SVR shown in the next page. Based on the training data showing the comparison of forecasting performance of 3 models using validation data. It is obtained that the GA-SVR method achieved the best results. Therefore, this method was utilized to test the number of cycles by repeating for 10, 50, 100 and 500 iterations, successively. A radial basis function was applied for a kernel function, and for optimal parameters in the GA method, RMSE is occupied by calculating the fitness function and determining populations with 50 sizes.

In Table 3, the three parameters (C, ε, γ) are values of the GA-SVR process iteration is utilized to forecast the optimal cycling distance. In Figure 4, (a), (b), (c) and (d) are shown for each of 10, 50, 100, and 500 iterations, respectively. Further, GA-SVR algorithms have been implemented into a computer simulation code written in R programming.

The three parameters (C, ε, γ) were obtained at each iteration. Afterwards, these parameters were added into the SVR model. The operating time frame for the GA-SVR method calculation is shown in Table 4, tests with validation data, in seconds and classified into some of 10, 50, 100, and 500 iterations, with the result of 0.35, 1.03, 2.00 and 10.02 minutes respectively. Remarks are gained that as the number of computations iterated increases, the computation time increases with the GA-SVR method. For comparing between 500 iterations and 100 iterations, we found that 500 iterations have given a lower error value with 10.02 minutes, while 100 iterations have taken time only 2 minutes. Consequently, it can be concluded that 100 iterations are the most effective. The RMSE and MAPE values of the model are calculated as seen on Table 4.

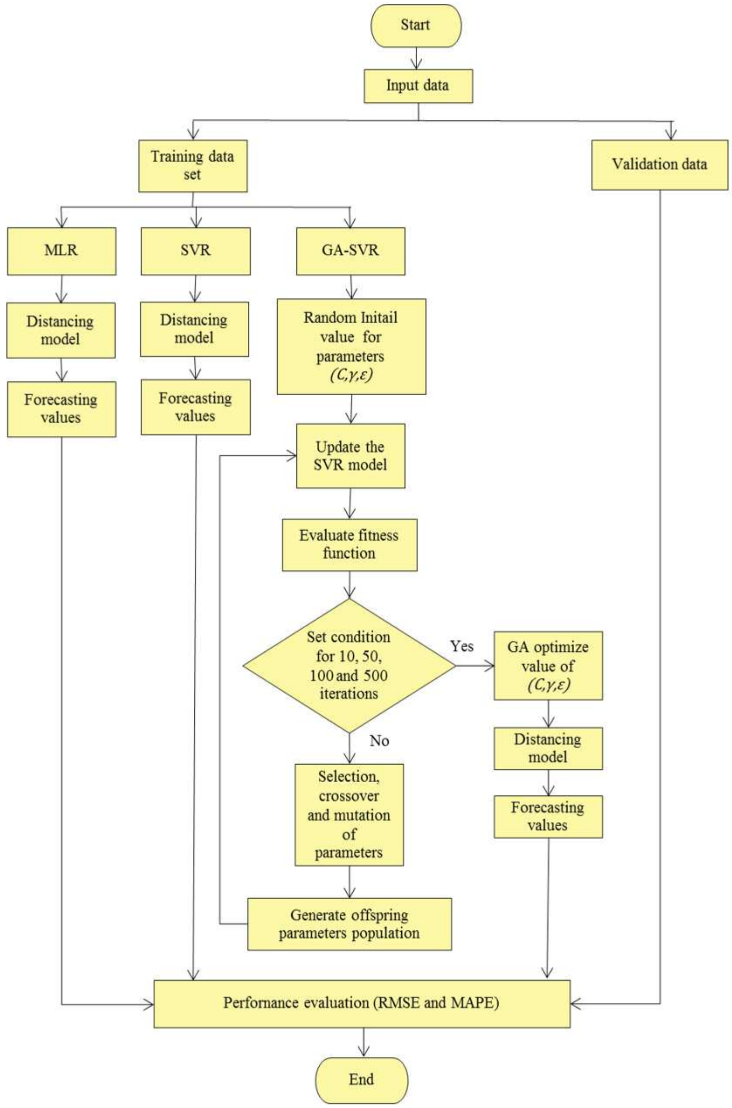

Figure 2: The modeling process for forecasting cycling distance

4 Conclusions

The data is gathered from the cycling distance of each person in Ladkrabang, Bangkok by collecting information with the questionnaire. The forecast model is created using three different techniques, namely the MLR, SVR, and GA-SVR models. The best and most effective model is selected by measuring the tolerances of RMSE and MAPE. In the GA-SVR section, GA is needed to find the optimal three parameters (C, ε, γ), insert them into the SVR to improve and replicate them for 10, 50, 100 and 500 iterations and measure the error to evaluate the best value. The results with 500 iterations yielded the lowest error, RMSE is 1.314 and MAPE is 0.091, but 500 iterations took the program's calculation time to 10.02 minutes, compared to 100 repetitions which gave the program computation time for only 2 minutes, but with similar tolerances. Therefore, the iteration of 100 times is sufficient as well as it will not fill in a variety of times to calculate. The forecasting cycling distance model mentioned above is

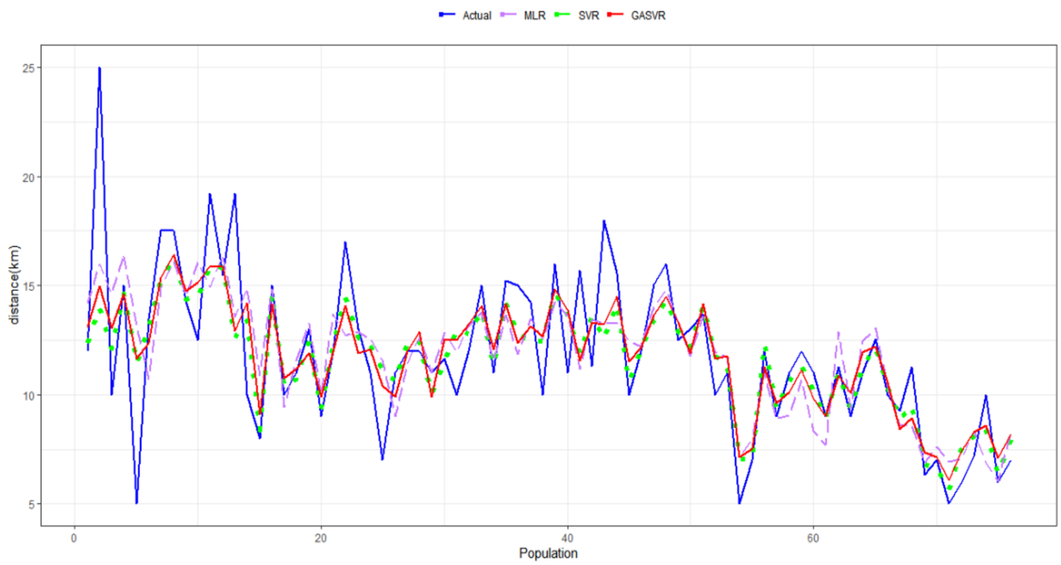

Figure 3: Comparisons between actual and forecast

Table 3: Iteration comparisons of GA-SVR parameters

Iterations	C	ε	γ
10	1.516	0.110	0.297
50	1.211	0.099	0.290
100	1.294	0.078	0.047
500	1.891	0.049	0.023

Table 4: Comparisons of the efficiency of GA-SVR

Method	Time (minutes)	RMSE	MAPE
GA-SVR			
10 iteration	0.35	1.529	0.113
50 iteration	1.03	1.501	0.108
100 iteration	2.00	1.389	0.094
500 iteration	10.02	1.314	0.091

intended to forecast the cycling distance of a person with physical limitations or the elderly by taking a sample for testing and can help cyclists set limits on how many kilometers each ride should be cycled in 30 minutes. The model also reduces the risks of falling ill for people with physical health limitations and helps to promote good health and well-being as well. The reader is hoping to develop calculating the cycling distance of 30 minutes by adding to the

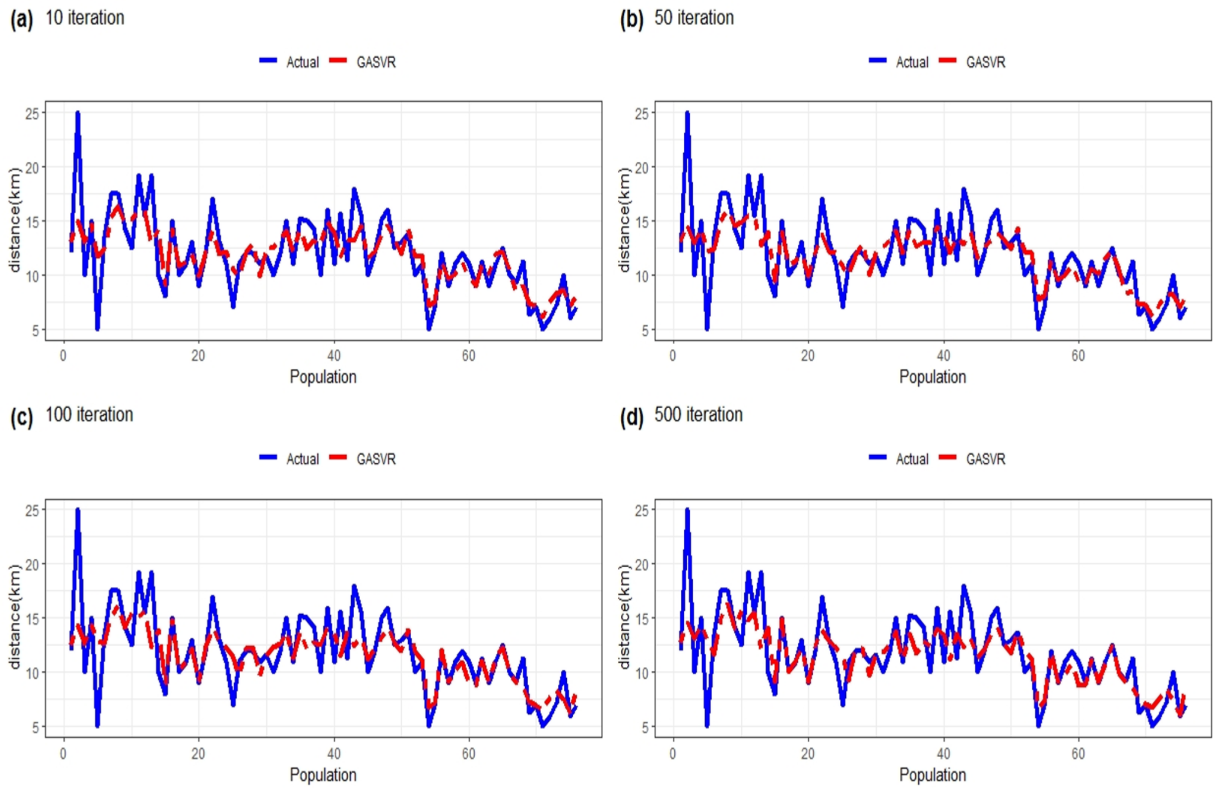

Figure 4: Plots between actual and forecast data sets

display screen of the stationary bike in the gym for further research. The results would be even better if we could forecast the suitable distance for each person.

References

[1] Buttar, H.S., Li, T. and Ravi, N. Prevention of cardiovascular diseases: Role of exercise, dietary interventions, obesity, and smoking cessation. *Experimental and Clinical Cardiology*. 2005. 10(4): 229–249.

[2] Cronkleton, E. *11 Benefits of Cycling, Plus Safety Tips*. Available:https://www.healthline.com/health/fitness-exercise/cycling-benefits, 2020.

[3] Mandal, A. Just 30 minutes of exercise each day can lower heart disease risk. Online. Available: shorturl.at/cqvLQ.

[4] Garber, C.E., Blissmer, B., Deschenes, M.R., Franklin, B., Lamonte, M.J., Lee, I.M., Nieman, D.C. and Swain, D.P. Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults. *Medicine & Science in Sports & Exercise*. 2011. 43(7): 1334-1359.

[5] Research and Markets. *Fitness Equipment: COVID-19*. Available: https://www.researchandmarkets.com

[6] Zander, A., Passmore, E., Mason, C. and Rissel, C. Joy, Exercise, Enjoyment, Getting out: A Qualitative Study of Older People's Experience of Cycling in Sydney, Australia. Journal of Environmental and Public Health. 2013. 2013(5): 453–547.

[7] Titze, S., Geus, B.D., Krenn, P., Bauman, A., Oja, P., Nash, B.R. and Kohlberger, T. Health bene?ts of cycling: A systematic review. *Scandinavian Journal of Medicine and Science in Sports*. 2011. 21(4): 496-509.

[8] Chuentawat, R., Kerdprasop, K. and Kerdprasop, N. The Demand Forecasting of Electrical Power Distribution Units by Support Vector Regression with 3 Folds Cross-Validation. *Journal of science and technology, Ubon Ratchathani University*. 2017. 19(1).

[9] Ogcu, G., Demirel, O.F., and Zaim, S. Forecasting Electricity Consumption with Neural Networks and Support Vector Regression. Procedia - Social and Behavioral Sciences. 2012. 58: 1576-1585.

[10] Xu, L., Zhang, W., Zhao, J., Cao, Z., Xie, R., Liu, X. and Hu, J. Support-vector-regression-based prediction of water holdup in horizontal oil-water flow by using a bicircular conductance probe array. Flow Measurement and Instrumentation. 2017. 57: 64-72.

[11] Yangming, G., Lu, Z., Xiaobin, C., Congbao, R., Zhengjun, Z. and Jiezhong, M. Time Series Adaptive Online Prediction Method Combined with Modified LS-SVR and AG. *Mathematical Problems in Engineering*. 2012.

[12] Suzukia, Y., Ibayashia, H., Kanedab, Y. and Minenoa, H. Proposal to sliding window-based support vector regression-KES2014. *18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems*. 2014. 35: 1615-1624.

[13] Yang, C.C. and Shieh, M.D. A support vector regression based prediction model of affective responses for product form design. *Computers & Industrial Engineering*. 2010, 59, 682–689

[14] Zangooei, M.H., Habibi, J. and Alizadehsani, R. Disease Diagnosis with a hybrid method SVR using NSGA- II. *Neurocomputing*. 2014. 136(14–29).

[15] Tang, X., Wang, L., Chang, J., Chen, J. and Sheng, V. S. Forecasting model based on information-granulated GA-SVR and ARIMA for producer price index. 2019.

[16] Guo, M. L., Li, D., Du, C., Jia, Z., Qin, X., Chen, L., Sheng, L. and Li, H. Prediction of the Busy Traffic in Holidays Based on GA-SVR. Advances in Computer Science and Information Engineering. 2012. 577-582.

[17] Zhou, J., Chang, H., Cheng, X. and Zhao, X. A Multiscale and High-Precision LSTM-GASVR Short-Term Traffic Flow Prediction Model. *Complexity*. 2020.

[18] Liu, S., Tai, H., Ding, Q., Li, D., Xu, L. and Wei, Y. A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction. Mathematical and Computer Modelling. 2013. 58(3–4): 458-46.

[19] Yotharak, P., Meeklueb, P. and Kongsang, S. *Mathematics Game with Health Development*. Bachelor's Thesis. King Mongkut's Institute of Technology Ladkrabang. 2018.

[20] Sanguansat, P. *Artificial Intelligence with Machine Learning*. Nonthaburi: IDC Premier. 2019.

[21] Awad, M. and Khanna, R. *Support Vector Machines for Classification*. 2015.

[22] Vapnik, V. *The Nature of Statistical Learning Theory*. New York: Springer. 1995.

[23] Kao, Y. S., Nawata, K. and Huang, C. Y. Predicting Primary Energy Consumption Using Hybrid ARIMA and GA-SVR Based on EEMD Decomposition. *Mathematics* 2020. 2020. 8(10).

[24] Parrella, F. Support Vector Machines for Regression. Online. Available: http://onlinesvr.altervista.org/.

[25] Granata, F., Gargano, R. and Marinis, G.D. Support Vector Regression for Rainfall-Runoff Modeling in Urban Drainage: A Comparison with the EPA's Storm Water Management Model. Water, 2016, 8(3), 69.

[26] Keerthi, S.S. and Lin, C.J. Asymptotic behaviors of support vector machines with Gaussian kernel. Neural Computation. 2003, 15(7): 1667-1689.

[27] Zhou, J., Chang, H., Cheng, X. and Zhao, X. A Multiscale and High-Precision LSTM-GASVR Short-Term Traffic Flow Prediction Model. *Complexity Problems Handled by Advanced Computer Simulation Technology in Smart Cities*. 2020.