Research Journal of Mathematical and Statistical Sciences ISSN 2320-6047
Vol. **8(1),** 1-5, January **(2020)** *Res. J. Mathematical and Statistical Sci.*

Comparative study of support vector regression model under different kernel functions

Kazeem Adewale Osuolale
Nigerian Institute of Medical Research, Lagos, Nigeria
whereisqosimadewale@gmail.com

Available online at: www.iscamaths.com , www.isca.in , www.isca.me
Received 5th June 2019, revised 20th October 2019, accepted 20th December 2019

Abstract

In spite of the fast growth in computer experiments, some computer models remain complex and often too time-consuming to be applied directly to mimic the experimentation conducted in the laboratory and predict the response of computer models using entirely a new set of data. To circumvent these problems and simplify the burden in using a simulator, a metamodel named as a support vector regression (SVR) model was used in this study as an emulator of a simple pendulum computer (SPC) model. The efficiencies of the SVR model under Gaussian Radial Basis (GRB) and B-Splines (BS) Kernel functions were examined. A SPC experiment developed using Orthogonal Array (OA)-based Latin Hypercube Design (LHD) was adopted to demonstrate the goodness of the metamodel. Further performance of SVR model at predicting the stoppage time of a pendulum bob was checked using the two kernel functions. Comparisons were also made on the performances of SVR model using R-square, relative average absolute error (RAAE) and relative maximum absolute error (RMAE), respectively. Estimated values for RAAE (0.0458) and RMAE (0.0500) provided by B-Splines functions are relatively smaller than the estimated values for RAAE (0.0604) and RMAE (0.2686) provided by the GRB function. Further results also showed that the SVR model with GRB kernel function has the estimated R^2(0.9936) that is relatively smaller than the one provided by the B-Splines function (0.9977). This study concluded that the SVR model with B-Splines kernel function outperforms the one that uses Gaussian Radial Basis function for modelling and predicting the stoppage time of a pendulum bob at untried inputs. The model developments and analyses were performed using 9.0.0.341360 (R2016a) MATLAB software.

Keywords: Computer experiment, modelling, predicting, stoppage time, simple pendulum computer model.

Introduction

Computer experiments are becoming more commonly employed in modern businesses, engineering and science due to their flexibility and wide applicability over the conventional physical experiments[1,2]. A computer experiment is carried out by using data obtained from a mathematical model, known as a computer model in lieu of the physical process[3]. Some circumstances may require that computer experiments be conducted before embarking on laboratory or physical experiments so that computer experiments can serve as a proxy for laboratory experiments. In a study, a researcher was quoted to have reported that the computer experiment was first performed at the Los Alamos Scientific Laboratory in 1953[4]. The OA (49, 3) LHD originally constructed by the author cited herein was used to develop a simple pendulum computer experiment and the SVR model is used as a metamodel to emulate a simple pendulum computer model[5]. The SVR algorithm was adopted for modelling and analysing simple pendulum computer experiments with the two kernel functions to ensure that the SVR efficiently captures the non-linearities in the simple pendulum model[6]. A simple pendulum experiment has been proposed as a novel application in the field of computer experiments by the researchers referenced in this study[7]. A simple pendulum model using a Gaussian process model as an emulator has been recently investigated[8]. An SVR model with B-Splines and SVR with Gausian Radial Basis kernel functions for modelling and predicting the stoppage time of simple pendulum computer experiments have been proposed separately and at different occasions by the authors under reference[9-10].

Methodology

The OA (49, 3) LHD adopted was scaled according to the assumed ranges for the input variables in the simple pendulum computer model using a standard method of scaling and this was subsequently used to develop a simple pendulum computer experiment using a simple pendulum computer model to produce the output of a computer experiment[11].

$$ML^2\frac{d^2\theta}{dt^2}+b\frac{d\theta}{dt}+MgLsin(\theta)=0$$

$$\frac{d^2\theta}{dt^2}=\frac{-b\frac{d\theta}{dt}-MgLsin(\theta)}{ML^2} \quad (1)$$

Where, $y=\frac{d^2\theta}{dt^2}$ = stoppage time of the bob, M = X_1 mass of the pendulum bob*(kg)*, L = X_2 length of the pendulum*(m)*, θ = X_3 displacement angle (degree), b = coefficient of friction (*Ns/rad*), g = acceleration due to gravity (ms^{-2}).

The coefficient of friction (*b*) is fixed at 0.02Ns/rad and the acceleration due to gravity (g) is 9.81ms^{-2}. The scaled input variables and the output simulated using Equation1constitute the output obtained for the training data sets as given in the result section. Then the SVR problem is formulated as the minimization of the following function:

$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\left(\phi_i + \phi_i^*\right) \quad (2)$$

where C is the penalty parameter and ε is the insensitive zone which can be chosen by the user. The parameter ε determines the width of the epsilon insensitive zone. The optimization problem in Equation (2) and its constraints can be written as a Lagrangian function:

$$L = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\left(\phi_i + \phi_i^*\right) - \sum_{i=1}^{n}\lambda_i(\varepsilon + \phi_i - y_i + (w.x_i) + b) - \sum_{i=1}^{n}\lambda_i^*(\varepsilon + \phi^*_i + y_i - (w.x_i) - b) - \sum_{i=1}^{n}\left(u_i\phi_i + u_i^*\phi_i^*\right) \quad (3)$$

where u_i and u_i^* are additional slack variables. Optimization problem is given in dual form as:

$$\max imize \begin{cases} -\frac{1}{2}\sum_{i,j=1}^{n}(\lambda_i - \lambda_i^*)(\lambda_j - \lambda_j^*)(x_i.x_j) \\ -\varepsilon\sum_{i=1}^{n}(\lambda_i + \lambda_i^*) + \sum_{i=1}^{n} y_i(\lambda_i - \lambda_i^*) \end{cases}$$

$$such\ that \begin{cases} \sum_{i=1}^{n}(\lambda_i - \lambda_i^*) = 0 \\ (\lambda_i - \lambda_i^*) \in [0, C] \end{cases} \quad (4)$$

and **w** is given as

$$w = \sum_{i=1}^{n}(\lambda_i^* - \lambda_i)x_i \quad (5)$$

The linear regression expressed is written as

$$f(x) = \sum_{i=1}^{n}(\lambda_i - \lambda_i^*)(x_i.x_i) + b \quad (6)$$

Where ($x_i.x_i$) is the dot products of two points and the variables λi, λ_i^* and b were estimated adopting the algorithm written by the authors cited in this study[12]. The authors developed a nonlinear regression model by replacing the dot product ($x_i.x_i$) with a kernel function K in the approximation function and equation (6) therefore gives the nonlinear approximation as follows:

$$f(x) = \sum_{i=1}^{n}(\lambda_i - \lambda_i^*)K(x_i.x_j) + b \quad (7)$$

An SVR technique has the input variables as the training inputs while the output is the training target. The algorithm adopted implements the model as:

$$\{N_{sv}, \beta, bias\} = f(Xdata(X), Ydata(Y), \ker nel(Ker), \mathrm{Cos}\,t(C), loss, insensitivity(\in)) \quad (8)$$

The length, displacement angle and mass of the pendulum were normalized to lessen the dimension effect of each variable and prevent an inconsistent prediction of the training target. Providing a better picture of metamodel efficacy also requires metrics for prediction accuracies. This is required for validation purpose and three different measures of model and prediction accuracies used were R^2, Relative Average Absolute Error (RAAE) and Relative Maximum Absolute Error (RMAE) as illustrated below:

RAAE

$$RAAE = \frac{\sum_{i=1}^{n}| y_i - \hat{y}_i|}{n.SD} \quad (9)$$

RMAE

$$RMAE = \frac{\max\left(| y_1 - \hat{y}_1|, | y_2 - \hat{y}_2|, ..., | y_n - \hat{y}_n|,\right)}{SD} \quad (10)$$

R-Square

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \quad (11)$$

According to this study, '*n*' is defined as the number of runs in the experiment, $\hat{y}_i$ is the predicted response estimate for the simulated value y_i, *SD* stands for standard deviation and $\bar{y}$ is the mean of the observed values. The lower the value of *RAAE and RMAE* estimated, the more accurate the SVR model and the larger the value of R^2 obtained, the more accurate the SVR model. It has been suggested that a good model would require $R_p^2 \geq 0.99^2$. It was also stated that the value of R_p^2 is greater than 0.9 for all computer models only when the sample size is greater than 40[13].

Results and discussion

The experimental data developed using the OALHD and the simple pendulum model is given in Table-1 as well as the results of the support vector regression (SVR) model with B-Splines and Gaussian Radial Basis Functions in Table-2.

The experiment was performed with a new set of data which is the 50% of the data used in the experiment in order to assess the prediction accuracy of the SVR model. These data used specific assumed ranges of design variables and used as the test data. The simulated test data were predicted and the plots of the predicted *y* against the simulated output were given in Figure-1 and Figure-2. The two functions were compared using the estimated values of R^2, RAAE and RMAE as shown in Table-3.

Table-1: Simple pendulum computer experimental data (training datasets).

Run	X_1	X_2	X_3	Y		X_1	X_2	X_3	Y
1	0.2270	8.4470	0.0119	8.8500	26	0.3499	27.8905	0.0123	5.4000
2	0.2319	12.8375	0.0129	8.2200	27	0.3548	32.2809	0.0133	4.6500
3	0.2368	17.2279	0.0139	7.5800	28	0.3597	36.6714	0.0144	3.8000
4	0.2418	21.6184	0.0150	6.9700	29	0.3647	10.9558	0.0166	2.3900
5	0.2467	26.0088	0.0160	6.3900	30	0.3696	15.3463	0.0176	2.9600
6	0.2516	30.3993	0.0170	5.8500	31	0.3745	19.7368	0.0186	3.5100
7	0.2565	34.7898	0.0181	5.3300	32	0.3794	24.1272	0.0125	4.4200
8	0.2614	9.0742	0.0130	6.9100	33	0.3843	28.5177	0.0135	3.3100
9	0.2663	13.4647	0.0141	6.4000	34	0.3892	32.9081	0.0145	2.8400
10	0.2713	17.8551	0.0151	5.8700	35	0.3942	37.2986	0.0155	3.2500
11	0.2762	22.2456	0.0161	5.3400	36	0.3991	11.5830	0.0178	2.6300
12	0.2811	26.6360	0.0172	4.8200	37	0.4040	15.9735	0.0188	3.3100
13	0.2860	31.0265	0.0182	4.2900	38	0.4089	20.3640	0.0126	2.4800
14	0.2909	35.4170	0.0120	7.3900	39	0.4138	24.7544	0.0136	2.6200
15	0.2958	9.7014	0.0142	5.3500	40	0.4187	29.1449	0.0147	3.1600
16	0.3008	14.0919	0.0153	4.8700	41	0.4237	33.5353	0.0157	3.7300
17	0.3057	18.4823	0.0163	4.3500	42	0.4286	37.9258	0.0167	3.5000
18	0.3106	22.8728	0.0173	3.7500	43	0.4335	12.2102	0.0189	2.4000
19	0.3155	27.2633	0.0183	2.6800	44	0.4384	16.6007	0.0128	2.5400
20	0.3204	31.6537	0.0122	6.3600	45	0.4433	20.9912	0.0138	3.1300
21	0.3253	36.0442	0.0132	5.6300	46	0.4482	25.3816	0.0148	3.7300
22	0.3303	10.3286	0.0154	3.8600	47	0.4532	29.7721	0.0158	4.3400
23	0.3352	14.7191	0.0164	3.1000	48	0.4581	34.1626	0.0169	4.1900
24	0.3401	19.1095	0.0175	2.7800	49	0.4630	38.5530	0.0179	4.8600
25	0.3450	23.5000	0.0185	3.1200					

Table-2: Results of SVR for the simple pendulum computer experiment.

Kernel Function Name	SV	Bias	Sum Beta	$\|w_0\|$^2
Gaussian RBF	38 (77.6%)	0.000	2.808404	28.174398
B-Splines	42 (85.7%)	0.000	0.030001	0.351313

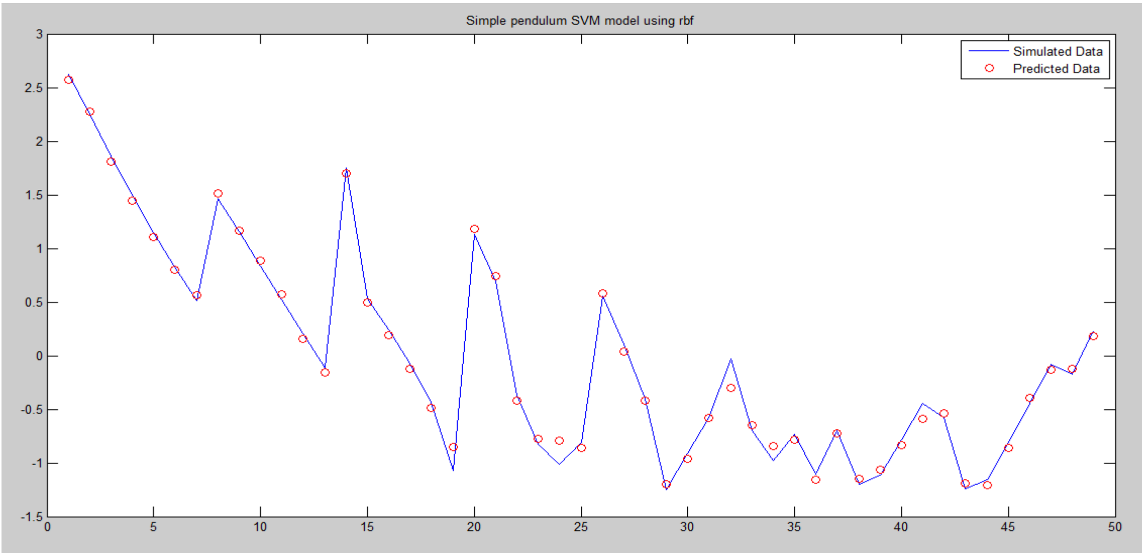

Figure-1: The Predicted *y* against Simulated Output using GRB Function.

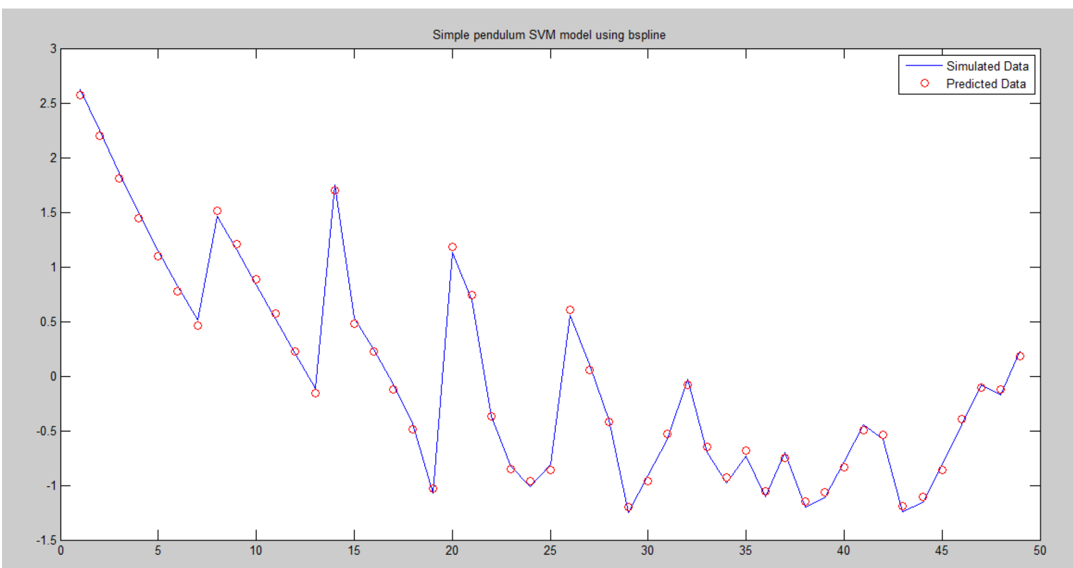

Figure-2: The Predicted *y* against Simulated Output using B-Splines Function.

Table-3: Comparison of results for the SVR model.

Kernel Function	R^2	RAAE	RMAE
Gauss. RBF	0.9936	0.0604	0.2686
B-Splines	0.9977	0.0458	0.0500

The SVR model trained with different number of Support Vectors as shown in Table-2 with only the Gaussian RBF that trained with 77.6% while the B-Spilnes trained with 85.7% of the experimental runs. Both Gaussian RBF and B-Splines kernel functions have zero bias. SVR with Gaussian RBF and B-Splines kernel functions indicated that not all the output from the experimental runs are support vectors. The loss and kernel functions were used to control SVR model and additional capacity control, C. The ε-insensitive loss function used fixed values of e=0.05 and C=8 having tested several values. Figure-1 and Figure-2 for SVR with GRB and B-Splines kernel functions displayed the plot of the predicted output against the simulated one. The model with the two kernel functions gave rough approximations of the test data. Predictions from the SVR with Gaussian RBF and B-Splines were not exact but close to the simulated output. SVR model did not interpolate the test data; that is, the predicted y did not equal the simulated output as shown in Figure-1 and Figure-2. The SVR model predicted the simulated output differently and is therefore regarded as an approximating metamodel. The estimated values of RAAE (0.0458) and RMAE (0.0500) provided by B-Splines functions are relatively smaller than the estimated values of RAAE (0.0604) and RMAE (0.2686) provided by the GRB function. Further results also showed that the SVR model with GRB kernel function has the estimated R^2 (0.9936) that is relatively smaller than the one provided by the B-Splines function (0.9977).

Conclusion

The estimated values of Relative Average Absolute Error (RAAE) and Relative Maximum Absolute Error (RMAE) provided by B-Splines functions are relatively smaller than the estimated values of RAAE and RMAE given by the GRB function. Further results also showed that the SVR model with GRB kernel function has the estimated R^2 that is relatively smaller than the one provided by the B-Splines function. This study therefore concludes that the support vector regression (SVR) model with B-Splines kernel function outperforms the one that uses Gaussian Radial Basis function for modelling and predicting the stoppage time of a pendulum bob at untried inputs.

References

1. Osuolale K.A., Yahya W.B. and Adeleke B.L. (2017). An enhanced algorithm for constructing optimal space-filling designs using hadamard matrices of orders 4λ and 8λ. *Annals, Computer Science Series*, 15(2), 9-17.

2. Osuolale K.A. and Yahya W.B. (2017). developing a javelin throw as a computer experiment using orthogonal array Latin hypercube design. Proceedings of the 1st International Conference of the Nigeria Statistical Society, Nigeria.

3. Osuolale K.A., Yahya W.B. and Adeleke B.L. (2014). Performing a simple pendulum experiment as a demonstrative example for computer experiments. *Annals. Computer Science Series*, 12, 33-38.

4. Strogatz S. (2003). The real scientific hero of 1953. New York Times. March 4, Editorial/Op-Ed.

5. Osuolale K.A., Yahya W.B. and Adeleke B.L. (2015). Construction of orthogonal array-based Latin hypercube designs for deterministic computer experiments. *Annals. Computer Science Series*, 13(1), 24-29.

6. Gunn S.R. (1998). Support vector machines for classification and regression. Technical Report, University of Southampton.

7. Osuolale K.A., Yahya W.B. and Adeleke B.L. (2014). Construction of space-filling designs for three input variables computer experiments. *WASET, International Journal of Computer, Control, Quantum and Information Engineering*, 8(9), 1624-1628.

8. Osuolale K.A., Yahya W.B. and Adeleke B.L. (2015). Modelling and analysis of simple pendulum computer experiment. Proceedings of the International Conference on Computer Science Research and Innovations (CoSRI), University of Ibadan.

9. Osuolale K.A.(2018). Modelling and analysis of simple pendulum computer experiment using a support vector regression model with B-Splines Kernel Function. Proceedings of the International Conference on Applied Physics and Mathematics, Dubai.

10. Osuolale K.A. (2019). Modelling and analysis of simple pendulum computer experiment using a support vector regression model. Thailand Statistician (in press).

11. MATLAB Basics (2003). Matlab/Simulink tutorial (2nd Ed.). School of Electrical, Electronic and Computer Engineering, University of Newcastle Upon Tyne.

12. Smola A.J. and Schölkopf B. (2004). A tutorial on support vector regression. *Statistics and computing*, 14(3), 199-222.

13. Sahama T. (2003). Some practical issues in the design and analysis of computer experiments (Ph.D. Thesis). Victoria University of Technology.