ENGINEERING JOURNAL

Article

Determination of Compressive Strength of Concrete by Statistical Learning Algorithms

Pijush Samui

Centre for Disaster Mitigation and Management, VIT University, Vellore, India
E-mail: pijush.phd@gmail.com

Abstract. This article adopts three statistical learning algorithms: support vector machine (SVM), lease square support vector machine (LSSVM), and relevance vector machine (RVM), for predicting compressive strength (f_c) of concrete. Fly ash replacement ratio (FA), silica fume replacement ratio (SF), total cementitious material (TCM), fine aggregate (ssa), coarse aggregate (ca), water content (W), high rate water reducing agent (HRWRA), and age of samples (AS) are used as input parameters of SVM, LSSVM and RVM. The output of SVM, LSSVM and RVM is f_c. This article gives equations for prediction of f_c of concrete. A comparative study has been carried out between the developed SVM, LSSVM, RVM and Artificial Neural Network (ANN). This article shows that the developed SVM, LSSVM and RVM models are practical tools for the prediction of f_c of concrete.

Keywords: Support vector machine, least square support vector machine, relevance vector machine, compressive strength, concrete.

ENGINEERING JOURNAL Volume 17 Issue 1
Received 18 July 2012
Accepted 16 September 2012
Published 1 January 2013
Online at http://www.engj.org/
DOI:10.4186/ej.2013.17.1.111

DOI:10.4186/ej.2013.17.1.111

1. Introduction

In the last years, a number of efficient statistical learning algorithms, e.g. support vector machine (SVM) [1, 2], least square support vector machine (LSSVM) [3], and relevance vector machine (RVM) [4] have been proposed. Successful applications of statistical learning algorithms have been reported for various fields [5-7]. This article adopts SVM, LSSVM and RVM for determination compressive strength (f_c) of concrete. SVM was developed by Vapnik and his coworkers in 1995, and it is based on the structural risk minimization (SRM) principle. LSSVM is proposed by taking with equality instead of inequality constraints to obtain a linear set of equations instead of a quadratic programming (QP) problem in the dual space [3, 8]. RVM is a sparse method for training generalized linear models [4]. It can be seen as probabilistic version of SVM. This study uses the database collected by Pala *et al.* [9]. Table 1 shows the dataset. The database contains information about fly ash replacement ratio (FA), silica fume replacement ratio (SF), total cementitious material (TCM), fine aggregate (ssa), coarse aggregate (ca), water content (W), high rate water reducing agent (HRWRA), age of samples (AS) and f_c. A comparative study has been carried out between the developed SVM, LSSVM and RVM models. The developed SVM, LSSVM and RVM provide equations for the prediction of f_c.

Table 1. Dataset used in this study.

FA (%)	SF (%)	TCM (kg/m^3)	ssa (kg/m^3)	ca (kg/m^3)	W (lt/m^3)	HRWRA (lt/m^3)	Age (days)	f_c (MPa)
0	0	500	724	1086	150	7.5	3	64.9
0	0	500	724	1086	150	7.5	7	75.5
0	0	500	724	1086	150	7.5	28	86.8
0	0	500	724	1086	150	7.5	56	87.2
0	0	500	724	1086	150	7.5	90	95.7
0	0	500	724	1086	150	7.5	180	97.7
15	0	500	700	1086	150	7.5	3	52.1
15	0	500	700	1086	150	7.5	7	66.4
15	0	500	700	1086	150	7.5	28	86
15	0	500	700	1086	150	7.5	56	94.8
15	0	500	700	1086	150	7.5	90	99.6
15	0	500	700	1086	150	7.5	180	106.3
25	0	500	683	1086	150	9.25	3	48
25	0	500	683	1086	150	9.25	7	65.7
25	0	500	683	1086	150	9.25	28	85.4
25	0	500	683	1086	150	9.25	56	90.4
25	0	500	683	1086	150	9.25	90	95.4
25	0	500	683	1086	150	9.25	180	107.8
45	0	500	650	1086	150	10.5	3	34.1
45	0	500	650	1086	150	10.5	7	49.2
45	0	500	650	1086	150	10.5	28	71.8
45	0	500	650	1086	150	10.5	56	85.4
45	0	500	650	1086	150	10.5	90	87.7
45	0	500	650	1086	150	10.5	180	97.7
55	0	500	634	1086	150	13	3	22.3
55	0	500	634	1086	150	13	28	57.4
55	0	500	634	1086	150	13	56	66.6
55	0	500	634	1086	150	13	90	72.8
55	0	500	634	1086	150	13	180	79.9
0	5	500	719	1086	150	8	3	58.3
0	5	500	719	1086	150	8	7	75.5
0	5	500	719	1086	150	8	28	87.8
0	5	500	719	1086	150	8	56	93.1
0	5	500	719	1086	150	8	90	93.6
0	5	500	719	1086	150	8	180	99.3
20	5	500	686	1086	150	9.25	3	46.3
20	5	500	686	1086	150	9.25	7	65.6
20	5	500	686	1086	150	9.25	28	78.5

20	5	500	686	1086	150	9.25	90	90.3
20	5	500	686	1086	150	9.25	180	95.9
40	5	500	654	1086	150	11	3	30.5
40	5	500	654	1086	150	11	7	48.6
40	5	500	654	1086	150	11	56	80
40	5	500	654	1086	150	11	90	83.4
40	5	500	654	1086	150	11	180	88.3
0	0	400	710	1157	160	4	3	35
0	0	400	710	1157	160	4	28	60.7
0	0	400	710	1157	160	4	56	67.1
0	0	400	710	1157	160	4	90	70.5
0	0	400	710	1157	160	4	180	70.6
15	0	400	690	1157	160	4.4	3	29.3
15	0	400	690	1157	160	4.4	28	56
15	0	400	690	1157	160	4.4	56	63.4
15	0	400	690	1157	160	4.4	90	68.5
15	0	400	690	1157	160	4.4	180	72.1
25	0	400	660	1157	160	4.8	3	24.7
25	0	400	660	1157	160	4.8	7	33.7
25	0	400	660	1157	160	4.8	28	49.3
25	0	400	660	1157	160	4.8	56	60.8
25	0	400	660	1157	160	4.8	90	66.2
25	0	400	660	1157	160	4.8	180	70.2
45	0	400	634	1157	160	5.2	3	14.5
45	0	400	634	1157	160	5.2	7	20.3
45	0	400	634	1157	160	5.2	28	43.9
45	0	400	634	1157	160	5.2	90	61.2
45	0	400	634	1157	160	5.2	180	63.7
55	0	400	621	1157	160	5.5	3	13.6
55	0	400	621	1157	160	5.5	7	19.8
55	0	400	621	1157	160	5.5	28	37.3
55	0	400	621	1157	160	5.5	56	47.1
55	0	400	621	1157	160	5.5	180	63.2
0	5	400	688	1157	160	5.5	3	37.3
0	5	400	688	1157	160	5.5	7	53
0	5	400	688	1157	160	5.5	28	69.4
0	5	400	688	1157	160	5.5	56	72.1
0	5	400	688	1157	160	5.5	180	74.5
20	5	400	662	1157	160	5.5	3	28.9
20	5	400	662	1157	160	5.5	7	42.1
20	5	400	662	1157	160	5.5	28	62.3
20	5	400	662	1157	160	5.5	56	69.9
20	5	400	662	1157	160	5.5	90	72.4
40	5	400	636	1157	160	6	3	14.5
40	5	400	636	1157	160	6	7	20.5
40	5	400	636	1157	160	6	28	44.6
40	5	400	636	1157	160	6	56	55.3
40	5	400	636	1157	160	6	90	59.1
40	5	400	636	1157	160	6	180	68.4
0	0	410	609	1132	205	0	3	26.1
0	0	410	609	1132	205	0	7	36.9
0	0	410	609	1132	205	0	28	50.8
0	0	410	609	1132	205	0	56	57.1
0	0	410	609	1132	205	0	90	58.1
0	0	410	609	1132	205	0	180	60.6
15	0	410	589	1132	205	0	3	23.3
15	0	410	589	1132	205	0	7	32.3
15	0	410	589	1132	205	0	28	48.9
15	0	410	589	1132	205	0	56	55.7
15	0	410	589	1132	205	0	90	62.6
15	0	410	589	1132	205	0	180	64.8

25	0	410	576	1132	205	0	3	18.4
25	0	410	576	1132	205	0	7	26.2
25	0	410	576	1132	205	0	28	41.7
25	0	410	576	1132	205	0	56	49.1
25	0	410	576	1132	205	0	90	53.7
25	0	410	576	1132	205	0	180	57.9
45	0	410	549	1132	205	0	3	13.4
45	0	410	549	1132	205	0	28	35.6
45	0	410	549	1132	205	0	56	47
45	0	410	549	1132	205	0	90	54.1
45	0	410	549	1132	205	0	180	56.6
55	0	410	536	1132	205	0	3	7.8
55	0	410	536	1132	205	0	7	11.3
55	0	410	536	1132	205	0	28	24
55	0	410	536	1132	205	0	56	33.7
55	0	410	536	1132	205	0	180	48.4
0	5	410	605	1132	205	0	3	27.4
0	5	410	605	1132	205	0	7	39.2
0	5	410	605	1132	205	0	28	57.3
0	5	410	605	1132	205	0	56	59.6
0	5	410	605	1132	205	0	90	67.3
0	5	410	605	1132	205	0	180	66.3
20	5	410	578	1132	205	0	3	20.1
20	5	410	578	1132	205	0	28	52.9
20	5	410	578	1132	205	0	56	60.7
20	5	410	578	1132	205	0	180	68
40	5	410	552	1132	205	0	3	11.4
40	5	410	552	1132	205	0	7	11.68
40	5	410	552	1132	205	0	28	38.7
40	5	410	552	1132	205	0	90	48.7
40	5	410	552	1132	205	0	180	58.4
20	5	410	578	1132	205	0	90	63.7
40	5	500	654	1086	150	11	28	71.1
45	0	400	634	1157	160	5.2	56	54.1
0	0	400	710	1157	160	4	7	48.4
15	0	400	690	1157	160	4.4	7	39.9
45	0	410	549	1132	205	0	7	18.4
55	0	500	634	1086	150	13	7	36.4
20	5	410	578	1132	205	0	7	30.6
55	0	400	621	1157	160	5.5	90	52.9
0	5	400	688	1157	160	5.5	90	73.7
40	5	410	552	1132	205	0	56	45.9
55	0	410	536	1132	205	0	90	41.4
20	5	400	662	1157	160	5.5	180	76
20	5	500	686	1086	150	9.25	56	85.8

2. Details of SVM

SVM uses the following expression for the prediction of output variable (*y*):

$$y = w.\phi(x) + b \tag{1}$$

where $\phi(x)$ expresses the high-dimensional feature space which is nonlinearly mapped from the input space x, b is bias and w is weight.

This article adopts FA, SF, TCM, ssa, ca, W, HRWRA, and AS as input variables. The output of SVM is f_c. Thus

$$x = [FA, SF, TCM, ssa, ca, HRWRA, AS]$$

and

$$y=[f_c].$$

The value of *w* and *b* have been estimated by minimizing the regularized risk function,

$$\text{Minimize:}\quad \frac{1}{2}\left\|w^2\right\|+C\frac{1}{N}\sum_{i=1}^{N}L_\varepsilon\left(y_i f\left(x_i\right)\right)$$

$$L_\varepsilon\left(y_i, f\left(x_i\right)\right)=\begin{cases}0 & \left|y_i-f\left(x_i\right)\leq\varepsilon\right|\\ \left|y_i-f\left(x_i\right)\right|-\varepsilon & \text{others}\end{cases} \tag{2}$$

where $L_\varepsilon\left(y_i, f\left(x_i\right)\right)$ is ε-insensitive loss function and ε is error insensitive zone. To minimize the effect of noise data, positive slack variables (ζ_i and ζ_i^*) have added in Eq. (2).

$$\begin{aligned}&\text{Minimize:}\quad \frac{1}{2}\left\|w^2\right\|+C\frac{1}{N}\sum_{i=1}^{N}\left(\zeta_i+\zeta_i^*\right)\\&\text{Subjected to:}\quad y_i-w.\phi\left(x_i\right)-b\leq\varepsilon+\zeta_i\\&\qquad w.\phi\left(x_i\right)+b-y_i\leq\varepsilon+\zeta_i^*\\&\qquad \zeta_i\geq 0\quad \zeta_i^*\geq 0\end{aligned} \tag{3}$$

By introducing kernel function $K\left(x_i, x_j\right)$, the above optimization problem (3) can be written in the following way [1]:

$$\begin{aligned}&\text{Maximize:}\quad \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\alpha_i-\alpha_i^*\right)\left(\alpha_j-\alpha_j^*\right)K\left(x_i,x_j\right)-\varepsilon\sum_{i=1}^{N}\left(\alpha_i-\alpha_i^*\right)+\sum_{i=1}^{N}y_i\left(\alpha_i-\alpha_i^*\right)\\&\text{Subject to:}\quad \sum_{i=1}^{N}\left(\alpha_i-\alpha_i^*\right)=0\\&\qquad \alpha_i,\alpha_i^*\in[0,C]\end{aligned} \tag{4}$$

where α_i, α_i^* are Lagrange multipliers. The final equation of SVM takes the following form:

$$y=\sum_{i=1}^{N}\left(\alpha_i-\alpha_i^*\right)K\left(x_i,x\right)+b \tag{5}$$

To develop the SVM, the data have been divided into the following groups:

- Training dataset: This is required to construct the SVM model. This article uses the same training dataset as used by Pala *et al.* [9].
- Testing Dataset: This is required to verify the developed SMV. This article uses the same testing dataset as used by Pala *et al.* [9].

This study adopts the radial basis function:

$$K\left(x,x_k\right)=\exp\left\{-\frac{\left(x_k-x\right)^T\left(x_k-x\right)}{2\sigma^2}\right\},\quad k, l=1,\ldots,N$$

DOI:10.4186/ej.2013.17.1.111

where σ is the width of radial basis function and T is transpose) as kernel function. The data have been normalized between 0 and 1. The program of SVM has been constructed by using MATLAB.

3. Details of RVM

RVM uses the following equation for the prediction of output *(y)*.

$$y = a'\phi(x) = \sum_{i=1}^{n} a_i K(x, x_i) + a_0 \tag{6}$$

where x is input, $K(x, x_i)$ is kernel function, n is number of data and a is weight. In this study,

$$x = [FA, SF, TCM, ssa, ca, HRWRA, AS]$$

and

$$y = [f_c]$$

The likelihood of the complete data set can be written as

$$p(y|a,\sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\|y - a\varphi\|^2\right\} \tag{7}$$

To prevent overfitting, automatic relevance detection (ARD) prior is set over the weights.

$$p(a/\alpha) = \prod_{i=0}^{n} \mathrm{N}(\alpha_i / 0, \alpha_i^{-1}) = \prod_{i=0}^{n} \frac{\alpha_i}{\sqrt{2\pi}} \exp\left(-\frac{(\alpha_i \alpha_j)^2}{2}\right) \tag{8}$$

where α is a hyperparameter vector that controls how far from zero each weight is allowed to deviate [10].

The posterior distribution over the weights is thus given by:

$$p(a/y, \alpha, \sigma^2) = \frac{p(y/a, \sigma^2) p(a/\alpha)}{p(y/\alpha, \sigma^2)} = (2\pi)^{-(n+1)/2} |\Sigma|^{-1/2} \exp\left[-\frac{1}{2}(a - \mu)' \Sigma^{-1} (a - \mu)\right] \tag{9}$$

where the posterior covariance and mean are respectively:

$$\Sigma = (\sigma^{-2}\Phi'\Phi + A)^{-1}$$
$$\mu = \sigma^{-2}\Sigma\Phi' y \tag{10}$$

For uniform hyperpriors over α and σ2, one needs only maximize the term $p(y/\alpha, \sigma^2)$:

$$p(y/\alpha, \sigma^2) = \int p(y/a, \sigma^2) p(a/\alpha)$$
$$= (2\pi)^{-n/2} |\sigma^2 I + \Phi A^{-1} \Phi'|^{-1/2} \times \exp\left[-\frac{1}{2} y' (\sigma^2 I + \Phi A^{-1} \Phi')^{-1} y\right] \tag{11}$$

Maximization of this quantity is known as the type II maximum likelihood method [11, 12] or the "evidence for hyper parameter" [13]. Hyper parameter estimation is carried out in iterative formulae, e.g., gradient descent on the objective function [14]. The outcome of this optimization is that many elements of this go to infinity such that w will have only a few nonzero weights that will be considered as relevant vectors.

This study adopts the same training dataset, testing dataset, kernel function and normalization technique for the RVM as used by the SVM. MATLAB has been used to develop RVM.

4. Details of LSSVM

LSSVM adopts the following equation for prediction of output (*y*)

$$y = w^T \varphi(x) + b \tag{12}$$

where *w* is weight, *b* is bias, *x* is input variable and $\varphi(x)$ is non-linear mapping function. In this study,

$$x = [FA, SF, TCM, ssa, ca, HRWRA, AS]$$

and

$$y = [f_c].$$

LSSVM uses the following optimization problem determination of *w* and *b*:

$$\text{Minimize: } \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^{N} e_k^2$$

$$\text{Subject to: } y(x) = w^T \varphi(x_k) + b + e_k \text{ , } k = 1, \ldots, N. \tag{13}$$

where e_k is the random errors and γ is a regularization parameter in determining the trade-off between minimizing the training errors and minimizing the model complexity.

The following equation has been obtained by solving the above optimization problem and it has been used for prediction of f_c [15, 16]:

$$f_c = y = \sum_{k=1}^{N} \alpha_k K(x, x_k) + b \tag{14}$$

where $K(x, x_k)$ is kernel function and α_k is lagrange multipliers.

This study adopts the same training dataset, testing dataset, kernel function and normalization technique for the LSSVM as used by the SVM and RVM. The program of LSSVM has been constructed by MATLAB.

5. Results and Discussion

For SVM, the design value of C, ε and σ have been determined by a trial and error approach. The design values of *C*, ε and σ are 100, 0.01 and 2 respectively. The best SVM produces 115 support vectors. The performance of training and testing dataset has been determined by using the design values of *C*, ε and σ.

Figure 1 shows the performance of training and testing for the SVM. This article employs coefficient of correlation (*R*) to assess the performance of SVM. For a good model, the value of *R* should be close to one. It is observed from Fig. 1 that the value of *R* is close to one for training as well as testing dataset. So, the developed SVM predicts f_c fairly well. The developed SVM presents the following equation (by substituting $K(x, x_k) = \exp\left\{ -\frac{(x_k - x)^T (x_k - x)}{2\sigma^2} \right\}$, $\sigma = 2$, $b = 0$ and $N = 130$ in Eq. (5) for the prediction of f_c):

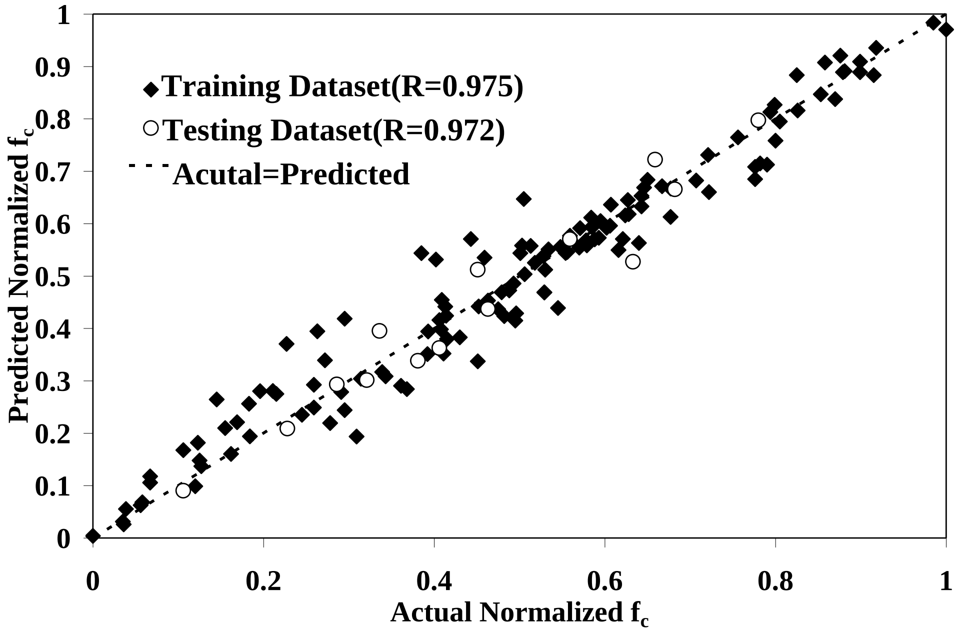

Fig. 1. Performance of the SVM.

$$f_c = \sum_{i=1}^{130} \left(\alpha_i - \alpha_i^*\right) \exp\left\{ - \frac{(x_k - x)^T (x_k - x)}{8} \right\} \tag{15}$$

The value of $\left(\alpha_i - \alpha_i^*\right)$ is shown in Fig. 2.

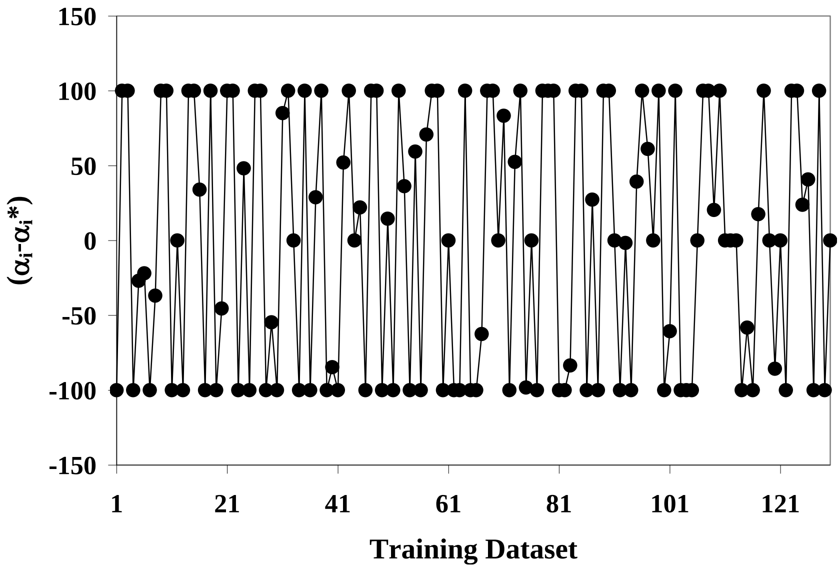

Fig. 2. Values of $\left(\alpha_i - \alpha_i^*\right)$ for the SVM.

In RVM, the trial and error approach has been adopted for determining the design value of σ. The developed RVM gives best performance at $\sigma = 1$. Therefore, the design value of σ is 1.

Figure 3 shows performance of RVM model. It is observed from Fig. 3 that the value of R is close to one for training as well as testing dataset. So, the developed RVM has the capability for predicting f_c. The developed RVM gives the following equation for prediction of f_c:

$$f_c = \sum_{i=1}^{130} a_i \exp\left\{ -\frac{(x_k - x)^T (x_k - x)}{2} \right\} \tag{16}$$

6. Conclusion

This study successfully applied SVM, RVM and LSSVM for the prediction of f_c of concrete. 130 datasets have been utilized to develop the models. User can use the developed equations for practical purposes. The developed RVM, SVM and LSSVM give almost the same performance. The obtained variance from the RVM can be used to determine uncertainty. SVM and RVM produce sparse solutions. In summary, it can be concluded that SVM, RVM and LSSVM can be examined for solving different problems in concrete.

References

[1] V. N. Vapnik, *The Nature of Statistical Learning Theory*, New York: Springer-Verlag, 1995.

[2] B. Schölkopf, *Support Vector Learning*, R. Oldenbourg Verlag,M¨unchen. Doktorarbeit, TU Berlin. Online at: http://www.kernel-machines.org.

[3] J. A. K. Suykens, and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Process. Lett.*, vol. 9, pp. 293-300, 1999.

[4] M. E. Tipping, "The relevance vector machine," *Advances in Neural Information Processing Systems*, vol. 12, pp. 652-658, 2000.

[5] D. Zhao, W. Zou, and G. Sun, "A fast image classification algorithm using support vector machine," *Proceeding 2nd International Conference on Computer Technology and Development*, pp. 385-388, 2010.

[6] H. Suetani, A. M. Ideta, and J. Morimoto, "Nonlinear structure of escape-times to falls for a passive dynamic walker on an irregular slope: Anomaly detection using multi-class support vector machine and latent state extraction by canonical correlation analysis," *Proceeding IEEE Int. Conference on Intelligent Robots and Sys.*, pp. 2715-2722, 2011.

[7] D. Matić, F. Kulić, M. Pineda-Sánchez, and I. Kamenko, "Support vector machine classifier for diagnosis in electrical machines: Application to broken bar," *Expert Syst. Appl.,* vol. 39, no. 10, pp. 8681-8689, 2012.

[8] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines*, Singapore: World Scientific, 2002.

[9] M. Pala, E. Özbay, A. Öztaş, and M. I. Yuce, "Appraisal of long-term effects of fly ash and silica fume on compressive strength of concrete by neural networks," *Constr. Build. Mater.*, vol. 21, no. 2, pp. 384-394, 2007.

[10] B. Schölkopf, and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, Cambridge: MIT Press, 2002.

[11] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis, 2nd ed.*, New York: Springer, 1985.

[12] G. A. Wahba, "Comparison of GCV and GML for choosing the smoothing parameters in the generalized spline-smoothing problem," *Ann. Statist.,* vol. 13, no. 4, pp. 1378-1402, 1985.

[13] D. J. MacKay, "Bayesian methods for adaptive models," *Ph.D. Thesis*, California Institute of Technology, 1992. Online at: http://resolver.caltech.edu/CaltechETD:etd-01042007-131447.

[14] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. of Machine Learning Research*, vol. 1, pp. 211-244, 2001.

[15] V. N. Vapnik, *Statistical Learning Theory*, New York: Wiley, 1998.

[16] A. J. Smola, and B. Schölkopf, "On a kernel based method for pattern recognition, regression, approximation and operator inversion," *Algorithmica*, vol. 22, pp. 211-231. 1998.