Computers and Structures 89 (2011) 1430–1439

Contents lists available at ScienceDirect

Computers and Structures

journal homepage: www.elsevier.com/locate/compstruc

Support vector regression based shear strength modelling of deep beams

Mahesh Pal *, Surinder Deswal

Department of Civil Engineering, NIT Kurukshetra, 136119 Haryana, India

A R T I C L E I N F O

Article history:
Received 27 November 2010
Accepted 17 March 2011
Available online 7 April 2011

Keywords:
Support vector machines
Deep beam
Shear strength prediction
Strut-and-tie method
Back-propagation neural network

A B S T R A C T

Support vector regression based modelling approach was used to predict the shear strength of reinforced and prestressed concrete deep beams. To compare its performance, a back-propagation neural network and the three empirical relations was used with reinforced concrete deep beams. For prestressed deep beams, one empirical relation was used. Results suggest an improved performance by the SVR in terms of prediction capabilities in comparison to the empirical relations and back propagation neural network. Parametric studies with SVR suggest the importance of concrete cylinder strength and ratio of shear span to effective depth of beam on strength prediction of deep beams.

1. Introduction

Various modelling approaches are being used to predict the behaviour of structures and their components by civil engineers. The traditional modelling approaches are based on empirical relationships derived using experimental data. Number of models has been proposed to predict the compressive strength and shear strength of beams and columns depending on different experimental conditions and assumptions. Several studies suggest the data specific nature of these models. As the design of a structure or a structural component may requires an iterative process in which the assumed model behaviour converges with the experimental behaviour, thus requiring a cost efficient computational technique.

Within last decade, researchers have explored the potential of back-propagation artificial neural networks (ANN) to solve various civil engineering problems. In structural engineering, neural networks have successfully been applied to several areas such as structural analysis and design [1–3], structural damage assessment [4–6], prediction of compressive strength of concrete mixes [7–18], shear strength prediction of reinforced concrete beams [19–24] and compressive strength of columns [25–28] as well as in non destructive strength assessment [29].

Determination of suitable architecture and various user-defined parameters has been a major issue in the design of an ANN [30]. ANN based modelling algorithm requires setting up of different learning parameters (like learning rate, momentum), the optimal number of nodes in the hidden layer and the number of hidden layers. In most of the reported applications, selection of number of hidden layers and the nodes in hidden layer is done by using a rule of thumb or trying several arbitrary architectures to select one that gives the best performance with test dataset. A suitable value of parameters like learning rate and momentum is also required for selected hidden layers and nodes. Design of a back-propagation neural network also involves in using a non-linear optimisation problem that may results in a local minima. During training process a large number of training iterations may force ANN to over train, which may affect the predicting capabilities of the model. Several studies suggested using a validation dataset (i.e. a dataset other than the training dataset) to have an idea about the suitable number of iterations for a specific dataset. This may be a problem for studies where number of dataset is limited, like one of concrete strength prediction. Recent studies [31–33] suggest the usefulness of genetic algorithm to find the optimal architecture of ANN.

An alternative modelling technique, called Support Vector Machines [34], has recently been applied to the field of civil engineering and provides improved performance in comparison to empirical relations and back-propagation neural network [35–42]. Keeping in view the better performance by the support vector machines, present study examines its potential in predicting the shear strength of reinforced concrete deep beams and prestressed deep beams. The results obtained by the support vector machines were compared with the [43–44] codes and strut-and-tie methods. The results were also compared with a back-propagation neural network to highlight the efficiency of the proposed method.

2. Support Vector Regression (SVR)

Support vector machines are classification and regression methods, which have been derived from statistical learning theory [45]. The Support vector machines based classification methods is based

* Corresponding author. Tel.: +91 1744 238353.
E-mail address: mpce_pal@yahoo.co.uk (M. Pal).

0045-7949/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compstruc.2011.03.005

Nomenclature

a	shear span of deep beam
A_S	area of non-prestressed longitudinal steel reinforcement
b_w	web width of deep beams
C	regularization parameter
D	effective depth of beam
D	kernel specific parameters for polynomial kernel
h	overall depth of prestressed deep beam
K	kernel function
L	effective span of reinforced deep beams
fyh	yield strength of horizontal reinforcement
fyv	yield strength of vertical web reinforcement
f_c	concrete cylinder strength of reinforced deep beams
F_y	yield strength of non-prestressed longitudinal steel reinforcement
F_{pe}	effective prestressing force after both immediate and long-term losses
F_c	compressive force in concrete strut of prestressed deep beams
V	shear strength (in kN)
ρ_h	ratio of horizontal web reinforcement
ρ_s	ratio of longitudinal reinforcement to area of concrete
ρ_v	ratio of vertical web reinforcement
γ	kernel specific parameter for RBF kernel

on the principle of optimal separation of classes. If the classes are separable – this method selects, from among the infinite number of linear classifiers, the one that minimise the generalisation error, or at least an upper bound on this error, derived from structural risk minimisation. Thus, the selected hyper plane will be one that leaves the maximum margin between the two classes, where margin is defined as the sum of the distances of the hyper plane from the closest point of the two classes [34].

Vapnik [34] proposed ε-Support Vector Regression (SVR) by introducing an alternative ε-insensitive loss function. This loss function allows the concept of margin to be used for regression problems. The purpose of the SVR is to find a function having at most ε deviation from the actual target vectors for all given training data and have to be as flat as possible [46]. For a given training data with k number of samples be represented by $\{\mathbf{x}_i, y_i\}$, $i = 1, \ldots, k$, where $\mathbf{x}_i$ is input vector and y_i is the target value, a linear decision function can be represented by

$$f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + b \tag{1}$$

where $\mathbf{w} \in \mathbf{R}^N$ and $b \in \mathbf{R}$. $\langle \mathbf{w}, \mathbf{x} \rangle$ represents the dot product in space $\mathbf{R}^N$. In Eq. (1), vector $\mathbf{w}$ determine the orientation of a discriminating plane whereas scalar b determine the offset of the discriminating plane from the origin. A smaller value of $\mathbf{w}$ indicates the flatness of Eq. (1), which can be achieved by minimising the Euclidean norm defined by $\|\mathbf{w}\|^2$ [34]. Thus, an optimisation problem for regression can be written as [46]:

$$\text{minimise} \quad \frac{1}{2}\|\mathbf{w}\|^2$$

$$\text{subject to} \begin{cases} y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - b \leqslant \varepsilon \\ \langle \mathbf{w}, \mathbf{x}_i \rangle + b - y_i \leqslant \varepsilon \end{cases} \tag{2}$$

The optimisation problem in Eq. (2) is based on the assumption that there exists a function that provides an error on all training pairs which is less than ε. In real life problems, there may be a situation like one defined for classification by [47]. So, to allow some more error, slack variables ξ, ξ' can be introduced and the optimisation problem defined in Eq. (2) can be written as below to deal with infeasible constraints of the optimization problem (2) [46]:

$$\text{Minimise} \quad \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{k}(\xi_i + \xi'_i)$$

$$\begin{aligned} &\text{Subject to } y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - b \leqslant \varepsilon + \xi_i \\ &\langle \mathbf{w}, \mathbf{x}_i \rangle + b - y_i \leqslant \varepsilon + \xi'_i \end{aligned} \tag{3}$$

and $\xi_i, \xi'_i \geqslant 0$ for all $i = 1, 2, \ldots, k$.

The constant $C > 0$ is a user-defined parameter which determines the trade-off between the flatness of the function and the amount by which the deviations to the error more than ε can be tolerated. The minimization problem in Eq. (3) is called the primal objective function. It was found that that in most cases the optimization problem defined by Eq. (3) can easily be solved by converting it into a dual formulation [47]. The optimisation problem in Eq. (3) can be solved by replacing the inequalities with a simpler form determined by transforming the problem to a dual space representation using Lagrangian multipliers [48].

The Lagrangian of Eq. (3) can be formed by introducing positive Lagrange multipliers $\lambda_i, \lambda'_i, \eta_i, \eta'_i$, $i = 1, \ldots, k$ and multiplying the constraint equations by these multipliers, and finally subtracting the results from the objective function (i.e, $\|\mathbf{w}\|^2$). The Lagrangian for Eq. (3) can now be written as:

$$\begin{aligned} L = {} & \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{k}(\xi_i + \xi'_i) - \sum_{i=1}^{k}\lambda_i(\varepsilon + \xi_i - y_i + \langle \mathbf{w}, \mathbf{x}_i \rangle + b) \\ & - \sum_{i=1}^{k}\lambda'_i(\varepsilon + \xi'_i + y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - b) - \sum_{i=1}^{k}(\eta_i\xi_i + \eta'_i\xi'_i) \end{aligned} \tag{4}$$

The dual variables in Eq. (4) have to satisfy$\lambda_i, \lambda'_i, \eta_i, \eta'_i \geqslant 0$. The solution of the optimisation problem involved in the design of SVR can be obtained by locating the saddle point of the Lagrange function defined in Eq. (4). The saddle points of Eq. (4) can be obtained by equating partial derivative of L with respect to $\mathbf{w}$, b, ξ_i and ξ'_i to zero and getting:

$$\partial_w \mathbf{L} = \mathbf{w} - \sum_{i=1}^{k}(\lambda'_i - \lambda_i) \cdot \mathbf{x}_i = 0 \tag{5}$$

$$\partial_b \mathbf{L} = \sum_{i=1}^{k}(\lambda'_i - \lambda_i) = 0 \tag{6}$$

$$\partial_{\xi_i} \mathbf{L} = C - \lambda_i - \eta_i = 0 \tag{7}$$

$$\partial_{\xi'_i} \mathbf{L} = C - \eta'_i - \lambda'_i = 0 \tag{8}$$

Substituting Eqs. (5)–(8) in Eq. (4), results in the optimisation problem of maximizing:

$$-\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k}(\lambda'_i - \lambda_i)(\lambda'_j - \lambda_j)(\mathbf{x}_i \cdot \mathbf{x}_j) - \varepsilon\sum_{i=1}^{k}(\lambda'_i + \lambda_i) + \sum_{i=1}^{k}y_i(\lambda'_i - \lambda_i) \tag{9}$$

$$\text{subject to } \sum_{i=1}^{k}(\lambda'_i - \lambda_i) = 0 \text{ and } \lambda_i, \lambda'_i \in [0, C]$$

Dual variables η_i, η'_i are eliminated by using conditions in Eqs. (7) and (8) and can now be written as $\lambda'_i = C - \eta'_i$ and $\lambda_i = C - \eta_i$,