CONCRETE STRENGTH ESTIMATION USING RELEVANCE VECTOR MACHINES

JALE TEZCAN, QIANG CHENG
and ARIF CEKIC
Department of Civil and Environmental Engineering
Southern Illinois University
1230 Lincoln Dr., 62901, Carbondale, IL
United States
e-mail: jale@siu.edu

Department of Computer Science
Southern Illinois University Carbondale
1000 Faner Dr., 62901, Carbondale, IL
United States

Professional Service Industries, Inc.
45749 Helm Street
Plymouth, Michigan 48170
United States

Abstract

Concrete is a highly versatile structural material that has been used for many centuries. Recently, there has been a marked increase in the use of supplementary cementitious materials and chemical admixtures. This increase, while facilitating the adjustment of concrete properties for specific uses, has also complicated the task of estimating the relevant performance measures from the mix ingredients. For structural engineering purposes, the most relevant parameter is the 28-day compressive strength. An ideal strength estimation method should not only allow the use of different sets of predictive variables, but

Keywords and phrases: concrete compressive strength, mix design, relevance vector machines.

Received May 26, 2010

© 2010 Scientific Advances Publishers
also account for prediction uncertainty. This paper proposes a probabilistic approach for the estimation of the 28-day compressive strength of concrete, based on the state-of-the-art relevance vector machine (RVM). An RVM belongs to the class of sparse kernel classifiers, which are powerful tools in classification and regression. Recently, the support vector machine (SVM), a sparse kernel model with a functional form identical to that of the RVM, has already proved successful in modelling concrete behavior. The RVM-based approach proposed in this paper offers several advantages over the ones based on SVM. First, by using a probabilistic kernel, an RVM provides information about prediction uncertainty. Second, compared to an SVM, an RVM uses fewer kernel functions for comparable generalization performance, providing a sparser representation. Third, the RVM model parameters are automatically determined from the training set, unlike the SVM algorithm, where the selection of model parameters typically involves cross-validation. A demonstrative application comparing the two approaches is presented. The results from this study confirm the advantages of the proposed method and demonstrate its effectiveness.

1. Introduction

Concrete is one of the oldest construction materials. In its simplest form, concrete is a mixture of cement, water, and aggregates. With the introduction of supplementary cementitious materials and chemical admixtures, concrete has undergone a significant transformation over time. Due to its high versatility and adaptability, concrete is one of the most commonly used structural materials in today’s construction industry.

Compressive strength of concrete is the most relevant performance measure for structural design purposes. Mechanical properties such as the flexural strength, tensile strength, and modulus of elasticity can be estimated from the compressive strength. In general practice, characterization and acceptance of concrete is based on the results of the compressive strength tests performed on standard cylindrical specimens. The compressive strength is determined by dividing the failure load by the cross sectional area resisting the load, assuming uniform stress distribution. The average of at least two measurements from cylinders with a diameter of 6 inches and a height of 12 inches, made from the same mix is reported as a single test result (ASTM_C_39 [2]). To satisfy
the strength requirements of a particular design project, the average of three consecutive tests should equal or exceed the specified strength with an acceptable confidence level (ACI_318 [1]).

Assuming that the samples are prepared and tested properly, the compressive strength is mainly determined by the mix design, and in particular, the water-to-cementitious materials ratio. Mix proportions to achieve certain strength with reasonable confidence can be determined by repeated trials. However, this approach is inconvenient, time consuming, and costly. The number of trial mixtures can be minimized with the help of predictive models.

Recently, the use of sparse kernel classifiers in regression problems has proved very successful. One such classifier, support vector machine (SVM) has already been used in modelling the compressive strength (Gupta [4]), elastic modulus (Yan and Shi [21]), and fire damage in concrete specimens (Chen et al. [3]). In this paper, we propose the use of the relevance vector machine (RVM) for estimation of the compressive strength of concrete from the mix design, and demonstrate an example application comparing the two methods. The proposed approach, in addition, to providing information about prediction uncertainty, has computational advantages over the SVM-based methods.

The remainder of this paper is organized as follows. Section 2 provides a brief overview of the predictive models proposed by various researchers for concrete strength estimation from the mix ingredients. Section 3 describes the SVM and RVM regression algorithms and provides a flowchart of the proposed procedure. Section 4 demonstrates an example application comparing the SVM-to RVM-based approaches. Section 5 summarizes the findings of this study.

2. An Overview of the Predictive Models for Concrete Strength

The inverse relationship between the compressive strength and the water-to-cement ratio has been recognized since 1890s (Neville [13], Sear
Traditional empirical equations represent a specific set of mix ingredients and conditions. Modern concrete typically includes supplementary materials and additives such as fly ash, silica fume, furnace slag, air-entraining agents, and water reducers. The fact that concrete strength is not only a function of mix ingredients, but also the environmental factors complicates the development of a model for concrete strength. The ideal method for strength estimation should not only allow the use of different sets of predictive variables, but also account for prediction uncertainty. In this regard, nonparametric regression algorithms show tremendous potential because of their ability to model complex relationships without an assumed functional form. Many researchers developed nonparametric regression models for concrete strength by using different supervised machine learning architectures. The potential of artificial neural networks (ANNs) as concrete strength estimation tools has been investigated in many studies (Lai and Serra [10], Yeh [22], Hong-Guang and Ji-Zong [5], Kim and Kim [8], Kim et al. [7], Fazel Zarandi et al. [24]). Genetic algorithms were also used (Lim et al. [11], Yeh [23]). While ANNs and genetic algorithms are powerful regression and prediction tools, they are prone to performance degradation with noise and data outliers. While, this problem can be somewhat alleviated by adjusting the model architecture and relevant parameters, the existence of local minima or insufficient capability for adapting to noisy data limits the level of accuracy that the traditional machine learning algorithms can offer.

The support vector machine (SVM), a relatively new development, offers increased accuracy by always locating the global optimum (Vapnik [20]). Using kernel functions, an SVM transforms the training data into a high dimensional feature space, where complex nonlinear relationships can be modelled using linear algorithms. A subgroup of the training set is selected as “support vectors”, and a predictive model is formed by using the selected set. The use of support vectors not only removes the requirement of selecting the model size (the number of the hidden neurons in the ANNs), but also yields a sparse representation.
Although the SVM is a state-of-the-art technique, it has a number of limitations (Tipping [19]). Regarding the concrete strength estimation problem, the biggest shortcoming is that an SVM only predicts the expected (mean or median) values of the function and does not provide any information regarding the confidence levels. It is important to recognize that, while mix design is the main factor controlling the compressive strength, the fabrication, handling, curing, and testing procedures all contribute to the test result and deviation among the test results is expected. The knowledge of confidence levels, or equivalently an estimate of the population variance is very useful. In addition, the accuracy of the SVM depends on optimal selection of the model parameters, typically involving a cross validation procedure over a wide range of possible values.

3. Algorithm

3.1. Problem definition

Given a set of data points \( \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\} \) such that \( x_i \in \mathbb{R}^d \) and \( y_i \in \mathbb{R} \), estimate the unknown function \( f : \mathbb{R}^d \rightarrow \mathbb{R} \) describing the input-output mapping.

3.2. Support vector machine (SVM) regression

An SVM transforms the training data into a high dimensional feature space and locates a hyperplane, where the output can be written as a linear combination of the weighted input. Using the kernel representation (Smola and Schölkopf [17]), the unknown function can be written as:

$$ f(x) = \sum_{i=1}^{N} w_i K(x_i, x) + w_0, \quad (1) $$

where \( w_i \) and \( w_0 \) represent the model weights and the bias term, respectively, and \( K(x_i, x) \) is a suitable kernel function satisfying a set of
Radial basis function (rbf) is a commonly used kernel function:

\[ K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \quad \gamma > 0. \]  

(2)

The classical SVM algorithm ignores errors below a user-defined threshold (\(\varepsilon\)) through the use of the \(\varepsilon\)-insensitive loss function (Huber [6])

\[ |y - f(x)|_\varepsilon = \begin{cases} |y - f(x)| - \varepsilon, & \text{if} \ |y - f(x)| > \varepsilon, \\ 0, & \text{otherwise}, \end{cases} \]  

(3)

and solves the minimization problem:

\[
\begin{align*}
\min_{w, \delta, \delta^*} & \quad \frac{1}{2} w^T w + C \sum_{i=1}^{N} (\delta_i + \delta_i^*) \\
\text{subject to} & \quad y_i - f(x_i) \leq \delta_i + \varepsilon, \\
& \quad f(x_i) - y_i \leq \delta_i^* + \varepsilon, \\
& \quad \delta_i, \delta_i^* \geq 0, \ i = 1, \ldots, N.
\end{align*}
\]  

(4)

The parameter \(C\) in Equation (4) is a user-defined parameter controlling the trade-off between the model complexity and training errors. The slack variables \(\delta_i, \delta_i^*\), shown in Figure 1, are used to adjust the width of the tolerance zone as needed.
During regression, some of the weight coefficients are set to zero, effectively removing the corresponding training samples from the model. The predictive model is then expressed in terms of the remaining samples (the so-called support vectors). The number of support vectors depends on the specifics of the training set and the model parameters. If the number of support vectors is only a small fraction of the number of training examples, a highly sparse kernel representation is achieved.

One important limitation of the SVM is the lack of probabilistic outputs, only the expected (mean or median) values of the output are provided. In addition, the accuracy of the SVM algorithm depends on proper selection of the model parameters ($C$ and $\varepsilon$) in addition to a kernel type and its related parameters. For example, assuming an rbf kernel, there are three parameters ($C$, $\varepsilon$, and $\gamma$) to be specified by the user. The trade-off parameter $C$, has no intuitive meaning (Shawe-Taylor and Cristianini [15]). The optimal of value for the parameter $C$ can be determined over a wide range of values. A process called cross-validation is often used to assess the model performance. Optimal selection of the insensitivity parameter $\varepsilon$ requires the knowledge of the noise variance, which is generally unknown. Asymptotically, unbiased estimators for $\varepsilon$ have been proposed (Smola et al. [16], Kwok [9]), but they do not
consider sample size. The scale parameter $\gamma$ can be set either by using several trials with different values, or through a cross-validation procedure along with $C$ and $\varepsilon$. Considering that a three-parameter grid search with only 10 points per grid requires $10^3$ evaluations, the computational cost of ensuring global optimality may become high.

### 3.3. Relevance vector machine (RVM) regression

A relevance vector machine (Tipping [18]) is a probabilistic sparse kernel model that uses the same functional form as the SVM. Assume that the $(x, y)$ pairs in the training data are related through the same $f(x)$ given in Equation (1) with some added noise:

$$y_i = f(x_i) + n_i, \quad i = 1, \ldots, N,$$

where $n_i$ are independent samples from zero-mean Gaussian noise with variance $\sigma^2$. Therefore, the target values are assumed to have a Gaussian distribution with mean $f(x_i)$ and variance $\sigma^2$:

$$p(y_i|x) = \mathcal{N}(y_i|f(x_i), \sigma^2).$$

Assuming each observation is independent, the likelihood of the entire dataset can be written as

$$p(y|\sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} e^{-\frac{1}{2\sigma^2} \|y - \Phi w\|^2},$$

where $y = (y_1, \ldots, y_N)^T$, $w = (w_0, \ldots, w_N)^T$, and $\Phi_{N \times N+1}$ is a matrix $\Phi_{nm} = K(x_n, x_{m-1})$ and $\Phi_{n1} = 1$. To prevent overfitting, the weights are described by an automatic relevance determination Gaussian prior (Neal [12]):

$$p(w|\alpha) = \prod_{i=0}^{N} p(w_i|\alpha_i) = \prod_{i=0}^{N} \mathcal{N}(w_i|0, \frac{1}{\alpha_i}).$$
The use of individual $\alpha_i$ for each $w_i$ is what makes the RVM computationally efficient. To complete the probabilistic description of the weights, the inverse variances $\alpha_1, \ldots, \alpha_N$ and the noise variance $\sigma^2$ need to be defined. For consistency, let $\beta = 1 / \sigma^2$. Since both $\alpha$ and $\beta$ are scale parameters, they can be represented by a gamma distribution (Tipping [19]):

$$p(\alpha_i|a, b) = \frac{b^a \alpha_i^{a-1} e^{-b \alpha_i}}{\Gamma(a)}, \quad (9)$$

with $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$. The prior distribution over $\alpha$ is written as

$$p(\alpha) = \prod_{i=0}^N p(\alpha_i|a, b). \quad (10)$$

Similarly, $\beta$ is described by:

$$p(\beta|c, d) = \frac{d^c \beta^{c-1} e^{-d \beta}}{\Gamma(c)}. \quad (11)$$

The overall prior on $w$ can be evaluated by using marginalization over $\alpha$:

$$p(w|a, b) = \prod_{i=0}^N \int_0^\infty p(w_i|\alpha_i) p(\alpha_i|a, b) d\alpha_i, \quad (12)$$

or, equivalently,

$$p(w|a, b) = \prod_{i=0}^N \int_0^\infty \mathcal{N}(w_i|0, 1 / \alpha_i) \Gamma(\alpha_i|a, b) d\alpha_i. \quad (13)$$

Because gamma distribution is the conjugate prior to normal distribution, the above integral can be evaluated explicitly. The resulting distribution is a student’s $t$-distribution (Tipping [19]), which is highly peaked at the origin and yields a sparse representation for $\beta$. The hyperparameters $\alpha_0, \ldots, \alpha_N$ and $\beta$ are computed iteratively. In this Bayesian estimation framework, after computing the values of $\alpha_0, \ldots, \alpha_N$ and $\beta$, we obtain
the noise variance $\sigma^2$. The uncertainty from the input is automatically propagated through the hierarchies of the Bayesian model, defining the uncertainty of the output. In other words, we can automatically derive the confidence intervals regarding the estimated output values.

Figure 2 shows an outline of the RVM procedure for estimating concrete strength. The parameters of the training data set can be modified to fit other sets of ingredients.
4. Numerical Results

The RVM algorithm described in the previous section was trained and tested by using published data (Lim et al. [11]). Their data contains the slump and compressive strength measurements from a total of 189 sets of mixtures, 108 of which were proportioned to have a compressive strength between 40 and 80 MPa. The set of 108 samples is used in this study. Table 1 lists the mix parameters and their ranges. The list of ingredients for individual mixes can be obtained from the paper by Lim et al. [11].

Table 1. Mix proportion ranges

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_1$</td>
<td>Water-to-binder ratio</td>
<td>0.30-0.45</td>
</tr>
<tr>
<td>$p_2$</td>
<td>Water (kg/m³)</td>
<td>160-180</td>
</tr>
<tr>
<td>$p_3$</td>
<td>Fine-to-all aggregate weight ratio</td>
<td>0.37-0.53</td>
</tr>
<tr>
<td>$p_4$</td>
<td>Super plasticizer (kg/m³)</td>
<td>1.89-8.50</td>
</tr>
<tr>
<td>$p_5$</td>
<td>Fly ash-to-binder ratio</td>
<td>0.0-0.2</td>
</tr>
<tr>
<td>$p_6$</td>
<td>Air entraining agent (kg/m³)</td>
<td>0.036-0.078</td>
</tr>
</tbody>
</table>

Ten of the 108 samples were set aside for testing, and the remaining 98 were used in training the regression models. Therefore, the training set is composed of 98 input-output pairs $\{(x_1, y_1), (x_2, y_2), \ldots, (x_{98}, y_{98})\}$, where $x_i = [p_{1i}, p_{2i}, p_{3i}, p_{4i}, p_{5i}, p_{6i}]$ is a vector consisting of the six parameters defined in Table 1, and $y_i$ is the corresponding 28-day compressive strength.

4.1. Training of the SVM model

The SVM model was trained by using the radial basis kernel function. The learning parameters ($\varepsilon$, $C$) and the kernel parameter ($\gamma$) were determined through a five-fold cross validation procedure by using a grid search. Optimal parameters for the SVM were found to be $\varepsilon = 0.0128$, $C = 6.4$, and $\gamma = 0.18$. 
At the end of training, the SVM selected 75 of the 98 training samples as support vectors. Figure 3 shows the training residuals in terms of the number of standard errors. The dots in the figure show the training samples, and the circles represent the support vectors. The root-mean-square-error and the correlation coefficient for training were $\text{RMSE} = 1.1258$ and $r = 0.9929$, respectively.

![Training Residuals - SVM](image)

**Figure 3.** Training residuals and support vectors in SVM algorithm.

### 4.2. Training of the RVM model

The hyperparameters of the RVM algorithm are determined automatically, but the kernel function and its parameters need to be specified. Radial basis kernel was selected as in the SVM case. The recommended starting value for $\gamma$ is $(1/N)$, where $N$ is the number of training samples. Considering that there are 98 samples in the training set, the kernel parameter was set to $\gamma = 0.01$. This choice of $\gamma$ proved to be a suitable choice in terms of small prediction errors, and it was observed that the performance of the RVM was pretty stable around this value.
Figure 4 shows the training residuals corresponding to the training samples in addition to the “relevance vectors”. Only 7 of the 98 samples were selected as relevance vectors. The root-mean-square-error and the correlation coefficient for training were $\text{RMSE} = 1.4332$ and $r = 0.9885$, respectively. Note that the training residuals were reported in terms of the number of standard deviations, not standard errors as in the SVM case. This is because RVM is a probabilistic kernel model, and it is able to predict the standard deviation from the training set.

**Figure 4.** Training residuals and relevance vectors in RVM algorithm.

### 4.3. Comparison of the test results

The SVM and RVM models were tested using the 10 samples that were excluded from the training set. Figure 5 shows the predictions from the two models in addition to the measured compressive strengths of the test samples. As the figure shows, the predictions from both models are very close to the measurements. The root-mean-square-error (RMSE) and the correlation coefficient ($r$) from the SVM algorithm for testing are $\text{RMSE} = 1.171$ and $r = 0.9959$, respectively. For the RVM model, $\text{RMSE} = 1.048$ and $r = 0.9972$. 
5. Conclusion

This paper proposes an RVM-based regression algorithm for predicting the 28-day compressive strength of concrete in a probabilistic framework. An RVM is a probabilistic kernel model that uses the same functional form as the SVM. The key difference between the two is the ability of the RVM to predict the population variance in addition to the expected compressive strength. The RVM was trained and tested using published data, and its performance was compared to the SVM. A radial basis function was selected as the kernel function in both models. The parameters of the SVM were determined by using five-fold cross validation with grid search. The training residuals from the SVM were calculated in terms of the number of standard errors and plotted along with the support vectors. The training residuals from the RVM were calculated in terms of the number of standard deviations and plotted along with the relevance vectors. The RVM model employed dramatically fewer kernel functions, resulting in a much sparser model. Only 7 out of
98 samples were kept by the RVM algorithm, as opposed to 75 kept by the SVM. The two models were tested using the 10 samples that were set aside before training. The performance of the two models was evaluated using the total root-mean-square-errors, and the correlation coefficients from training and testing. It was observed that the generalization performance of the two models is comparable.

The following conclusions can be drawn from the results of this study:

(1) The proposed algorithm successfully predicts the 28-day compressive strength of concrete from the mix ingredients.

(2) The proposed RVM-based model provides the population variance in addition to the compressive strength, whereas the SVM only predicts the compressive strength.

(3) The RVM uses far fewer kernel functions compared to the SVM model for comparable generalization performance.

(4) The RVM model parameters are automatically determined from the training set, unlike the SVM algorithm, where the selection of model parameters typically involves cross-validation.

References

[1] ACI_318, Building Code Requirements for Structural Concrete and Commentary Detroit, American Concrete Institute 2008.


