Expert Systems with Applications 42 (2015) 885–893

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Modeling slump of ready mix concrete using genetic algorithms assisted training of Artificial Neural Networks

Vinay Chandwani *, Vinay Agrawal, Ravindra Nagar

Department of Civil Engineering, Malaviya National Institute of Technology Jaipur, Rajasthan, India

ARTICLE INFO

Article history:
Available online 6 September 2014

Keywords:
Artificial Neural Networks
Genetic algorithms
Back-propagation algorithm
Lavenberg Marquardt training algorithm
Concrete slump
Ready mix concrete

ABSTRACT

The paper explores the usefulness of hybridizing two distinct nature inspired computational intelligence techniques viz., Artificial Neural Networks (ANN) and Genetic Algorithms (GA) for modeling slump of Ready Mix Concrete (RMC) based on its design mix constituents viz., cement, fly ash, sand, coarse aggregates, admixture and water-binder ratio. The methodology utilizes the universal function approximation ability of ANN for imbibing the subtle relationships between the input and output variables and the stochastic search ability of GA for evolving the initial optimal weights and biases of the ANN to minimize the probability of neural network getting trapped at local minima and slowly converging to global optimum. The performance of hybrid model (ANN-GA) was compared with commonly used back-propagation neural network (BPNN) using six different statistical parameters. The study showed that by hybridizing ANN with GA, the convergence speed of ANN and its accuracy of prediction can be improved. The trained hybrid model can be used for predicting slump of concrete for a given concrete design mix in quick time without performing multiple trials with different design mix proportions.

1. Introduction

The mathematical relationships commonly used to describe the material behavior of concrete are available in the form of empirical formulae derived from experimental results. Although these empirical relationships in the form of regression equations are widely used and recommended for extracting knowledge about a particular property of concrete, yet these cannot be applied wherein the modeling problem involves a large number of independent variables or the interactions among the variables is either unknown or too complex to represent. In such cases the traditional technique of regression fails to yield the expected accuracy and predictability. Over the past few decades nature inspired computational tool, Artificial Neural Network (ANN) has been used for modeling the real world problems due to its immense ability to capture inter-relationships among input and output data pairs which are unknown, nonlinear or too difficult to formulate. This potential of ANN has been harnessed for wide applications in modeling the material behavior and properties of concrete. Notable among them are successful implementations in predicting and modeling compressive strength of self compacting concrete (Uysal & Tanyildizi, 2012), high performance concrete (Yeh, 1998), recycled aggregate concrete (Duan, Kou, & Poon, 2013), rubberized concrete (Abdollahzadeh, Masoudnia, & Aghababaei, 2011), fiber reinforced concrete (FRP)-confined concrete (Naderpour, Kheyroddin, & Ghodrati Amiri, 2010), durability of high performance concrete (Parichatprecha & Nimityonskul, 2009), predicting drying shrinkage of concrete (Bal & Buyle-Bodin, 2013), concrete mix design (Ji, Lin, & Lin, 2006) and prediction of elastic modulus of normal and high strength concrete (Demir, 2008).

One of the physical properties of concrete which plays an important role in the success of RMC industry is its workability. It signifies the ease, with which fresh concrete can be placed, compacted and finished at site with sufficient resistance to segregation. Being a quality assurance metric quantitatively measured as concrete slump value, it not only controls quality and uniformity of concrete from batch to batch but also acts a measure to ascertain the shelf life of the RMC during its transit course from manufacturing plant to subsequent placing at the construction site. Moreover it ensures that the RMC design mix is customized catering to the type of construction viz., heavily reinforced sections, lightly reinforced sections, road pavements, shallow sections or construction requiring intensive vibration, demanding high, medium, low, very low or extremely low workability concrete respectively. Recent applications of ANN modeling for concrete slump include prediction of slump and strength of ready mix concrete containing retarders and high strength concrete containing silica fume and plasticizers (Dias & Pooliyadda, 2001), predicting slump of fly ash

* Corresponding author.
E-mail addresses: chandwani2@yahoo.com (V. Chandwani), agrawal_vinay_2000@yahoo.com (V. Agrawal), ravindranagar@hotmail.com (R. Nagar).

http://dx.doi.org/10.1016/j.eswa.2014.08.048
0957-4174/

and slag concrete (Yeh, 2006), modeling slump of high strength concrete (Oztas et al., 2006), modeling slump of high performance concrete (Yeh, 2007) and modeling and analysis of concrete slump using laboratory test results (Jain, Jha, & Misra, 2008).

Back-propagation neural network (BPNN) due to its ability to map complex non-linear and unknown relationships is a preferred choice among researchers for modeling unstructured problems. BPNN is a multi-layer feed-forward neural network (MFNN) trained using back-propagation (BP) algorithm. The BP algorithm is a local search algorithm which employs gradient descent to iteratively update the weights and biases of the neural network, minimizing the performance function commonly measured in terms of a squared error between the actual and ANN predicted output. Despite its popularity as a universal function approximator and easy implementation, the BP algorithm is faced with inherent drawback of getting trapped in local minima and slow convergence. The reason for this drawback is attributed to random initialization of synaptic weights and biases prior to training a neural network. With every re-run of neural network during training phase, the BP algorithm evaluates a different set of final weights leading to trained neural network having different prediction performance and convergence speed. In order to minimize the BPNN's probability of inconsistency, it is necessary to develop an effective methodology for improving its prediction performance and convergence to global optima.

To overcome the inherent drawback of BP algorithm, genetic algorithms (GA) have been harnessed for evolving the optimal initial weights and biases for ANN. GA is a gradient free global optimization and search technique inspired by the evolutionary processes namely, natural selection and genetic variation, which allow simultaneous search for optimal solutions in different directions minimizing the chance of getting trapped in a local minimum and faster convergence. Successful implementations of this methodology can be found in Asadi, Shahrabi, Abbaszadeh, and Tabanmehr (2013), Irani and Nasimi (2011), Johari, Javadi, and Habibagahi (2011), Pendharkar (2009), Sedki, Ouazar, and El Mazoudi (2009), Tan, He, Nie, Zhang, and Hu (2014). Despite numerous applications of integrating GA with ANN in various fields of study, the methodology has not been explored so far for modeling slump of concrete. The study deals with amalgamating the universal function approximating ability of BPNN and the global search ability of GA for developing a robust computational tool for modeling slump of RMC.

The study has been organized into sections. Section 2 deals with data collection. Section 3 deals with the methodology, in which neural network modeling of concrete slump, its optimization using genetic algorithm assisted training and statistical performance measures have been discussed. Results, discussions and conclusions and future work have been dealt in Sections 4, 5, 6 respectively.

2. Data collection

The exemplar data for ANN were collected from the same RMC plant to mitigate any chance of change caused in the slump data due to change in composition of concrete mix constituents. The data comprised of concrete design mix constituents consisting of 560 mix proportions namely, cement, fly ash, sand (as fine aggregate), coarse aggregate 20 mm, coarse aggregate 10 mm, admixture, water-binder ratio and corresponding slump value.

3. Methodology

For conducting the study, the Neural Network Toolbox and Global Optimization Toolbox included in the commercially available software MATLAB R2011b (Version 7.13.0.564) were used to implement the BPNN and GA respectively.

3.1. ANN modeling of concrete slump

3.1.1. Preparing training, validation and test data sets

ANN is an information processing paradigm inspired by the learning ability of human brain. ANN therefore requires exemplar patterns to establish the underlying relationships between the input–output data pairs. Moreover, it is also necessary to assess the predictive power of the trained ANN when presented with examples not included in the neural network training. To facilitate training and testing of the neural networks, the collected data were randomized and split into training, validation and test data-sets. 70% of the data were used for training purpose and the remaining 30% data were equally divided and set aside for validation and testing of the trained ANN. The training data-set was used for training the ANN, enabling it to learn the relationships between the input and output data-pairs by systematic updating of the neural network weights and biases using BP algorithm. During the training phase, there is a tendency of the neural network to over-fit or over-learn the exemplar patterns presented during the training phase. This leads to poor generalization of the network when subjected to unseen data. Validation data-set is indirectly used during the training of ANN to monitor the over-fitting of the neural network and to act as a guide to stop the training of the neural network when the validation error begins to rise. Testing of the neural network is done after completion of the training phase. The test data set used during the testing phase evaluates the prediction performance of the trained neural network.

Efficient training of ANN requires that all representative patterns included in the exemplar data, should form a part of the training data-set. Hence, to allow the training data-set extend to the edges of modeling domain, it was ensured that extreme values (maximum and minimum values) of each constituent of total data-set were included in training data-set. Moreover data division should also reflect that training, validation and test data set is representative of the same population. Therefore, three ways split of data was done in such a way that the statistical parameters of Training, Validation and Test data sets viz., maximum value, minimum value, mean and standard deviation of each constituent are marginally different from each other. Table 1 shows the statistical parameters of the data used for training, validation and testing.

3.1.2. Preprocessing of data

The input data and output data generally comprise of different identities either having no or minimum similarities. Preprocessing or normalization of data eliminates the possibility of neural network bias towards the different identities and scales down all the input and output data preferably in a bound range [0, 1] or [−1, 1]. Scaling of inputs to the range [−1, 1] greatly improves the learning speed, as these values fall in the region of sigmoid transfer function where the output is most sensitive to the variations of the input values (Alshihri, Azmy, & El-Bisy, 2009). Linear scaling in the range [−1, 1] has been used in present study having function

$$x_{norm} = \frac{2 * (x - x_{\min})}{(x_{\max} - x_{\min})} - 1 \tag{1}$$

where x_{norm} is the normalized value of the variable x, $x_{\max}$ and $x_{\min}$ are the minimum and maximum values of variable x respectively.

3.1.3. Neural network architecture and training parameters

The architecture of an ANN consists of a number of artificial neurons connected through weighted connections. The artificial