TẠP CHÍ KHOA HỌC & CÔNG NGHỆ ĐẠI HỌC DUY TÂN
DTU Journal of Science and Technology 02(33) (2019) 23-27

TRƯỜNG ĐẠI HỌC DUY TÂN
DUYTAN UNIVERSITY

Xây dựng mô hình mạng nơ-ron thần kinh nhân tạo dùng cho phân tích hồi quy trong ngành xây dựng với lập trình trên Visual C# .NET

A neural network program for solving Regression analysis problems in construction engineering developed in Visual C# .NET

Hoàng Nhật Đức[a,*], Nguyễn Duy Tân[b], Nguyễn Trần Mộng Thùy[c]
Nhat Duc Hoang, Duy Tan Nguyen, Mong Thuy Nguyen Tran

[a]Viện Nghiên cứu Phát triển Công nghệ Cao, Đại học Duy Tân, 03 Quang Trung, Đà Nẵng, Việt Nam
Institute of Research and Development, Duy Tan University, 03 Quang Trung, Danang, Vietnam
[b]Phòng Tài nguyên và Môi trường thành phố Rạch Giá, Kiên Giang
Department of Resource and Environment, Rach Gia city, Kien Giang
[c]Công ty CP Tư vấn Đầu tư Xây dựng Hiệp Lực, thành phố Rạch Giá, Kiên Giang
Hiep Luc Investment Consultation Joint-Stock Company, Rach Gia city, Kien Giang

(Ngày nhận bài: 25/12/2018, ngày phản biện xong: 05/01/2019, ngày chấp nhận đăng: 01/04/2019)

Tóm tắt

Trong kỹ thuật xây dựng, phân tích hồi quy là một công cụ quan trọng được sử dụng trong phân tích dữ liệu và mô hình hóa các quá trình phức tạp. Để đơn giản hóa việc sử dụng mạng nơ ron nhân tạo (ANN) trong phân tích hồi quy, nghiên cứu này đã phát triển một chương trình ANN trong Visual C#, được gọi là VS-ANN. Khả năng phân tích của chương trình VS-ANN đã được kiểm chứng bởi hai ứng dụng về dự đoán cường độ nén bê tông và ước tính độ sâu sạt lở tại chân cầu. Kết quả về giá trị tương quan tốt R^2 (> 0.9) cho thấy chương trình VS-ANN là một công cụ hữu ích để thực hiện phân tích hồi quy trong kỹ thuật xây dựng.

Từ khóa: Phân tích hồi quy, mạng lưới thần kinh nhân tạo, kỹ thuật xây dựng, ngôn ngữ C#, cường độ nén, độ sâu sạt lở.

Abstract

In construction engineering, regression analysis is a crucial tool used in data analysis and modeling complex phenomena. To ease the employment of artificial neural network (ANN) in regression analysis, this research has developed an ANN program in Visual C#, called VS-ANN. NET framework 4.6.1. The capability of the VS-ANN program has been confirmed by two case studies of concrete compressive strength prediction and scour depth estimation. Good value of R^2 (>0.9) shows that the newly developed VS-ANN program can be a useful tool for performing regression analysis in construction engineering.

Keywords: Regression analysis, artificial neural network, construction engineering; Visual C#, compressive strength, scour depth.

1. Introduction

Regression analysis involves the identification of the mathematical relationship between a set of predictors and a variable of interest. This mathematical relationship is then employed for various modeling tasks in construction engineering [1 - 5] as well as in other industries [6]. Recent advancements

Email: hoangnhatduc@dtu.edu.vn

in machine learning have created intelligent data analysis methods that are proven to be more capable than conventional statistical approaches. Hence, these modern data analysis tools including neural network, support vector machine, regression tree, etc. have drawn great attention of various scholars as well as practicing engineers.

In construction engineering, empirical data modeling is a crucial research field. It is because researched phenomena in this field of study are inherently complex and heavily relied on empirical experiments. Hence, machine learning is often employed to generate a model of a system that can be used to predict the novel cases that have not been observed [7]. Therefore, the introduction of machine learning based regression methods is necessary for both practical use and educational purposes in engineering courses.

This study focuses on Artificial Neural Network (ANN) based on regression analysis models since its effectiveness in nonlinear model has been well demonstrated and its mathematical elegance [8, 9]. As an Artificial Intelligent (AI) technique, the structure of ANN and its inference process is capable of adapting dynamically in response to the update of data [10]. Based on various experimental works, ANN has demonstrated to be capable of delivering satisfactory predictive performances [11 - 13]. Considering the fact that open software for ANN implementation is relatively limiting, this study aims at contributing a system based on ANN algorithm for solving regression analysis problems in construction engineering. The rest of the article is organized as follows: the second section describes the formulation of ANN; the modeling capability of the newly constructed program is demonstrated in the third section; concluding remarks of this study are stated in the final section.

2. Artificial neural network and its implementation in Visual C# .NET framework

ANN is a powerful machine learning based on data analysis algorithm which is a model of the real biological neural networks. This machine learning approach attempts to simulate the knowledge acquisition and inference processes occurred in the human brain [14]. The ANN has been widely used for solving nonlinear regression analysis problems. It has been well demonstrated that an ANN with one hidden layer can sufficiently approximate very complex nonlinear functions [15].

The learning task of regression analysis is to generalize a mapping function $f: X \in R^D \to Y \in R^1$ where D denotes the number of input attributes [14]. An ANN model typically consists of the input, hidden, and output layers. The ANN structure used for function estimation can be expressed as follows [14]:

$$f(X) = b_2 + W_2 \times (f_A(b_1 + W_1 \times X) \qquad (1)$$

where W_1 and W_2 are weight matrices of the hidden layer and the output layer, respectively; N is the number of neurons in the hidden layer; $b_1 = [b_{11}, b_{12,}..., b_{1N}]$ denotes a bias vector of the hidden layer; b_2 is a bias vector of the output layer; f_A denotes an activation function. Herein, the commonly employed activation function is log-sigmoid [14].

The weight matrices and the bias vectors of an ANN are learnt through a training process that use the error backpropagation algorithm [16]. In addition, the Mean Square Error (MSE) is often used as the objective function to be minimized:

$$MSE = \min_{W_1, W_2, b_1, b_2} \frac{1}{M} \sum_{i=1}^{M} e_i^2 \qquad (2)$$

where M is the number of data samples; e_i represents an output error. $e_i = Y_{i,P} - Y_{i,A}$ ($Y_{i,P}$ and $Y_{i,A}$ denote predicted and actual outputs, respectively).

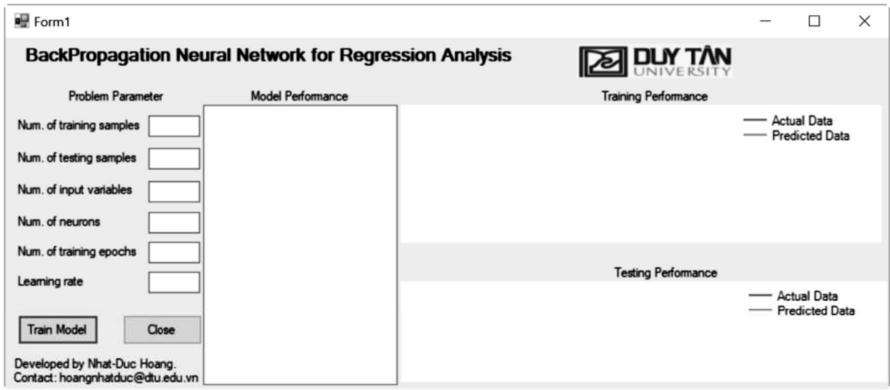


Fig. 1. The program interface

To ease the implementation of ANN used for regression analysis, a program called VS-ANN has been developed in Visual C# .NET (Framework 4.6.1). The graphical user interface (GUI) of the program is illustrated in Fig. 1. The program GUI is designed to simplify the application of ANN in construction engineering and still provides the users with selections to configure an ANN model. The users need to provide the input parameters including the number of training and testing samples, the number of input variables (or the predictors), the number of neurons used in the hidden layer, the number of training epochs, and the learning rate parameters. The training and testing data sets are provided in two separated ".csv" files: "TrainDataSet.csv" and "TestDataSet.csv". The default locations of these two files are "D://TrainDataSet.csv" and "D://TestDataSet.csv". Before the data is splitted, it is strongly recommended that the original dataset has been normalized using Z score equation:

$$Y_N = (Y_O - M_Y)/ S_Y \quad (3)$$

where Y_O and Y_N are the original and the normalized variables. M_Y and S_Y denote the mean and the standard deviation of the variable.

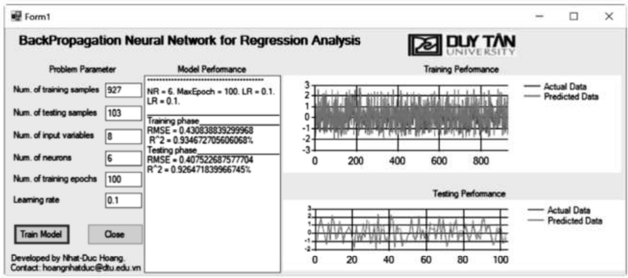
The original range of the output variable can be converted from its normalized value using the following equation:

$$Y_O = Y_N x S_Y + M_Y \quad (4)$$

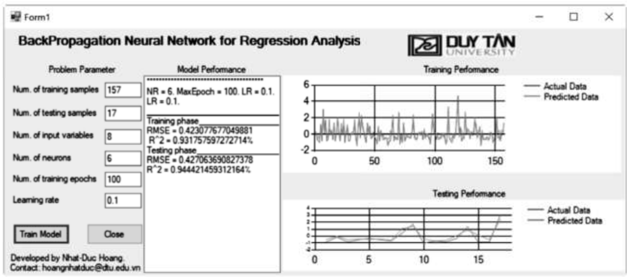
3. Software program applications

In this section, the VS-ANN program is applied in the two case studies: prediction of compressive strength of high performance concrete and prediction of scour depth of bridge piers. The original data set has been normalized using Z score equation and randomly divided into a training set (90%) and a testing set (10%). The number of neurons in the hidden layer is roughly set to be $2D_X/3 + 1$ which D_X is the number of input variables [17].

In construction engineering, high-performance concrete (HPC) has been widely employed in high-rise building/infrastructure projects for its superior strength, durability, and workability which exceed those of normal concrete [4].Since the compressive strength is widely used as the main criterion in defining the required quality of concrete, prediction of compressive strength can be very helpful in concrete mix design. This study employs 1030 data samples collected in [18] and can be assessed via [19]. The cement quantity (kg/m^3), blast furnace slag quantity (kg/m^3), fly ash quantity, water quantity, superplasticizer quantity, coarse aggregate quantity (kg/m^3), fine aggregate quantity (kg/m^3), concrete age (day) are employed as predictors. Concrete compressive strength (Mpa) is used as the dependent or output variable. The prediction results are shown in Fig. 2a. The model performance in the testing phase is demonstrated in terms of root mean square error (RMSE) and coefficient of determination (R^2): RMSE = 0.43 and R^2 = 0.92 which are a good predictive result. This means that 92% of the variation of the compressive strength can be explained by the VS-ANN program.



(a)



(b)

Fig. 2. Prediction results: (a) Case study 1 and (b) Case study 2

In the second application, the program is used to estimate the scour depth at bridge piers. Bridge scour is generally known as the removal of sediment (e.g. sand and gravel) from around bridge abutments or piers [20]. Scour which is caused by swiftly moving water can scoop out scour holes; this leads to the deterioration of the integrity of a bridge structure [21, 22]. Thus, models that can accurately estimate the scour depth at bridge piers are highly desirable. This study employs the data set including 8 variables that consider aspects of pier geometry, flow property, and river bed material.

The data set including 174 samples has been documented in [23]. The 8 predictors, including the flow depth y, the pier width perpendicular to the flow direction b_c, the pile-cap width b_{pc}, the longitudinal extension of pile cap face out from pier face L_u, the soil covering height Y, the ratio of the mean velocity to the critical velocity of sediment movement V/V_c, the median grain size d_{50}, and the river bed material geometric standard deviation σ_g, are employed to estimate the scour depth d_s of complex pier foundations [24]. The prediction performance of the VS-ANN program is reported in Fig. 2b. The model performance in the testing phase is as follows: RMSE = 0.41 and R^2 = 0.94 which demonstrates a high correlation of the observed and predicted values of the scour depth. The R^2 = 0.94 means that 94% of the variation of the scour depth can be explained by the VS-ANN program.

4. Conclusion

Regression analysis is an important data analysis tool widely used in construction engineering. To facilitate the employment of ANN in regression analysis, this study has developed a neural network program in Visual C#. NET framework. A simple GUI has been used to ease the implementation of the ANN algorithm. The effectiveness of the constructed program has been demonstrated via two case studies of compressive strength prediction and scour depth estimation. Good value of R^2 shows that the VS-ANN program can be a useful tool for performing function approximation problems in construction engineering.

References

[1] P. Chetchotisak, P. Ruengpim, D. Chetchotsak, S. Yindeesuk, Punching Shear Strengths of RC Slab-Column Connections: Prediction and Reliability, KSCE J Civ Eng, 22 (2018) 3066.

[2] Z. Zhongya, J. Xiaoguang, Prediction of Peak Velocity of Blasting Vibration Based on Artificial Neural Network Optimized by Dimensionality Reduction of FA-MIV, Mathematical Problems in Engineering, 2018 (2018) 12.

[3] N.-D. Hoang, A.-D. Pham, Q.-L. Nguyen, Q.-N. Pham, Estimating Compressive Strength of High Performance Concrete with Gaussian Process Regression Model, Advances in Civil Engineering, (2016) 8.

[4] A.-D. Pham, N.-D. Hoang, Q.-T. Nguyen, Predicting Compressive Strength of High-Performance Concrete Using Metaheuristic-Optimized Least Squares Support Vector Regression, Journal of Computing in Civil Engineering, 30 (2016) 06015002.

[5] P.S. Mangat, B.T. Molloy, Prediction of long term chloride concentration in concrete, Materials and Structures, 27 (1994) 338.

[6] W. Mendenhall, T.T. Sincich A Second Course in Statistics: Regression Analysis (7th Edition), Pearson2011.

[7] A. Çevik, A.E. Kurtoğlu, M. Bilgehan, M.E. Gülşan, H.M. Albegmprli, Support vector machines in structural engineering: a review, Journal of Civil Engineering and Management, 21 (2015) 261-281.

[8] P.-T. Ngo, N.-D. Hoang, B. Pradhan, Q. Nguyen, X. Tran, Q. Nguyen, V. Nguyen, P. Samui, D. Tien Bui, A Novel Hybrid Swarm Optimized Multilayer Neural Network for Spatial Prediction of Flash Floods in Tropical Areas Using Sentinel-1 SAR Imagery and Geospatial Data, Sensors, 18 (2018) 3704.

[9] N.-D. Hoang, An Artificial Intelligence Method for Asphalt Pavement Pothole Detection Using Least Squares Support Vector Machine and Neural Network with Steerable Filter-Based Feature Extraction, Advances in Civil Engineering, (2018) 1-12.

[10] M.-Y. Cheng, N.-D. Hoang, Groutability prediction of microfine cement based soil improvement using evolutionary LS-SVM inference model, Journal of Civil Engineering and Management, 20 (2014) 839-848.

[11] H. Tanyildizi, Prediction of the Strength Properties of Carbon Fiber-Reinforced Lightweight Concrete Exposed to the High Temperature Using Artificial Neural Network and Support Vector Machine, Advances in Civil Engineering, 2018 (2018) 10.

[12] L. Tang, C. Xiong, L. Zhang, Spatial Transferability of Neural Network Models in Travel Demand Modeling, Journal of Computing in Civil Engineering, 32 (2018) 04018010.

[13] H. Moayedi, S. Hayati, Applicability of a CPT-Based Neural Network Solution in Predicting Load-Settlement Responses of Bored Pile, International Journal of Geomechanics, 18 (2018) 06018009.

[14] T.-H. Tran, N.-D. Hoang, Predicting Colonization Growth of Algae on Mortar Surface with Artificial Neural Network, Journal of Computing in Civil Engineering, 30 (2016) 04016030.

[15] C.M. Bishop, Pattern Recognition and Machine Learning (Information Science and Statistics) Springer (April 6, 2011), ISBN-10: 03873107382011.

[16] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature, 323 (1986) 533-536.

[17] J. Heaton, Artificial Intelligence for Humans, Volume 3 Deep Learning and Neural Networks, Heaton Research, Inc., United States2015.

[18] I.C. Yeh, Modeling of strength of high-performance concrete using artificial neural networks, Cement and Concrete Research, 28 (1998) 1797-1808.

[19] I.-C. Yeh Concrete Compressive Strength Data Set, the UC Irvine Machine Learning Repository, http://archive.ics.uci.edu/ml/datasets/concrete+compressive+strength, (2007).

[20] D.S. Mueller, C.R. Wagner, Field Observations and Evaluations of Streambed Scour at Bridges, Office of Engineering Research and Development Federal Highway Administration, McLean, VA, (2005).

[21] L.P. Warren, Scour at Bridges: Stream Stability and Scour Assessment at Bridges in Massachusetts, U.S. Geological Survey, (2011).

[22] A.N. Kallias, B. Imam, Probabilistic assessment of local scour in bridge piers under changing environmental conditions, Structure and Infrastructure Engineering, 12 (2016) 1228-1241.

[23] K.-W. Liao, Y. Muto, J.-Y. Lin, Scour Depth Evaluation of a Bridge with a Complex Pier Foundation, KSCE Journal of Civil Engineering, 22 (2018) 2241-2255.

[24] N.-D. Hoang, K.-W. Liao, X.-L. Tran, Estimation of scour depth at bridges with complex pier foundations using support vector regression integrated with feature selection, Journal of Civil Structural Health Monitoring, (2018).