Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 127 (2017) 284–290

Energy Procedia

www.elsevier.com/locate/procedia

International Youth Nuclear Congress 2016, IYNC2016, 24-30 July 2016, Hangzhou, China

Prediction of components degradation using support vector regression with optimized parameters

Chunzhen Yang[a], Jingquan Liu[a,*], Yuyun Zeng[a], Guangyao Xie[a]

[a]*Department of Engineering Physics, Tsinghua University, Beijing, 100084, P. R. China*

Abstract

To solve the problem that traditional data-driven methods cannot acquire high accuracy with few monitoring data in nuclear power component degradation prediction, this article proposed the support vector regression (SVR) algorithm and provided a hybrid parameters optimization strategy for SVR using grid search and cross validation. This article analyzed the performance of SVR algorithm in three different cases: function fitting, multivariate regression and prediction of nuclear power plant pipeline corrosion. Results show that the performance of proposed SVR algorithm is better than classical BP neutral network by comparing prediction mean square error and squared correlation coefficient of two methods. The advantage of SVR algorithm in small-sample learning is also demonstrated and it turns out that the SVR algorithm is an effective approach for modelling component degradation with rare inspection data.

Peer-review under responsibility of the organizing committee of IYNC2016

Keywords: degradation prediction; support vector regression; parameter optimization; grid search; cross validation

1. Introduction

Degradation process of large scale structure or component is rather complex and modeling by mechanism will be difficult or even impossible in some occasion. Generally, data-driven methods are employed to model the degradation process and evaluate its condition [1]. However, failure data or condition monitoring data are very rare for many critical components in nuclear power plant because of their high reliability and therefore it brings some new challenges

* Corresponding author. Tel.: +86-106-279-4863
E-mail address: jingquan@tsinghua.edu.cn

1876-6102
Peer-review under responsibility of the organizing committee of IYNC2016
10.1016/j.egypro.2017.08.109

to data-driven modeling method. To solve the problem that traditional data-driven methods cannot acquire high accuracy in nuclear component degradation prediction with rare monitoring data, this article proposed a support vector regression (SVR) algorithm with optimized parameters to model and evaluate component degradation. Before applying SVR algorithm to the case of nuclear power plant pipeline wall thinning, this article demonstrated its application in function fitting and multivariate regression and the performance comparison was also conducted between SVR and classical BP neural network.

The paper is organized as follows. Section 2 briefly reviews the background and basic theory of SVR algorithm. The hybrid parameter optimization strategy for SVR model training is introduced in Section 3. Section 4 demonstrates the SVR application with three different cases. Finally, the main findings are summarized in Section 5.

Nomenclature

C	penalty factor
EFPY	effective full power years
k(**x, x'**)	kernel function
L	training sample number
MSE	mean square error
MPG	miles per gallon
N	test sample number
RBF	radial basis function
R^2	squared correlation coefficient
SVM	support vector machine
SVR	support vector regression
w	vector of weights
y_o	observed wall thickness value
y_p	predicted wall thickness value
γ	radial basis function width coefficient
ε	insensitive loss factor
σ	standard deviation of noise
δ	normally distributed noise
α_i^*, α_i	Lagrange multipliers
ξ_i , ξ_i^*	slack variables

2. Support vector regression algorithm

Based on the statistical learning theory, the Support Vector algorithm had been developed for decades [2]. In 2000s, the support vector regression was successfully applied to regression and time series prediction [3]. As a machine learning algorithm, support vector regression has many exclusive advantages in solving problems with small sample and high dimensionality [4]. In consideration of these advantages, the support vector regression is chosen for modeling the component degradation process in this work.

Suppose $\{(x_1, y_1), ..., (x_L, y_L)\}$ are the training samples and $f(\mathbf{x})$ is the regression function. For ε-SVR, it tries to form a "tube" whose radius is ε around the y values and errors are computed only for points lying outside this tube [1,5]. Minimizing the norm of **w** in Eq. (1) ensures the fitness precision.

$$f(\mathbf{x}) = (\mathbf{w} \cdot \mathbf{x}) + b \tag{1}$$

To cope with the outliers in training samples, penalty factor C and slack variables ξ_i and ξ_i^* are introduced. The penalty factor C determines the trade-off between the flatness of $f(\mathbf{x})$ and the amount up to which deviations larger than ε are tolerated. Hence we get the formulation as