Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Reliability Engineering and System Safety 92 (2007) 423–432
RELIABILITY ENGINEERING & SYSTEM SAFETY
www.elsevier.com/locate/ress

Forecasting systems reliability based on support vector regression with genetic algorithms

Kuan-Yu Chen*

Department of Business Administration, Far East College, 49 Chung Hua RD, Hsin-Shih, Tainan 744, Taiwan

Received 23 August 2005; received in revised form 16 December 2005; accepted 24 December 2005
Available online 28 February 2006

Abstract

This study applies a novel neural-network technique, support vector regression (SVR), to forecast reliability in engine systems. The aim of this study is to examine the feasibility of SVR in systems reliability prediction by comparing it with the existing neural-network approaches and the autoregressive integrated moving average (ARIMA) model. To build an effective SVR model, SVR's parameters must be set carefully. This study proposes a novel approach, known as GA-SVR, which searches for SVR's optimal parameters using real-value genetic algorithms, and then adopts the optimal parameters to construct the SVR models. A real reliability data for 40 suits of turbochargers were employed as the data set. The experimental results demonstrate that SVR outperforms the existing neural-network approaches and the traditional ARIMA models based on the normalized root mean square error and mean absolute percentage error.

Keywords: Support vector regression; Neural networks; Genetic algorithms; ARIMA

1. Introduction

The ability of accurately predicting product reliability is invaluable for manufacturing companies. Reliability predictions are used for various purposes, such as production planning, maintenance planning, reliability assessment, dectecting faults in manufacturing processes, evaluating risks and liabilities, prediction warranty costs, evaluating replacement policies, and assessing design changes, vendors, materials, and manufacturing processes. As reliability modeling and prediction play an increasingly important role in assessing the performance of engineering systems, intensive studies have been carried out to ensure system reliability. Existing studies on systems reliability modeling were based on lifetime distribution models, parts counts and parts stress, fault tree analysis and Markov models. Each of these approaches has achieved some success in certain cases and has come to be widely used. However, most existing analytical systems reliability modeling processes depend on certain priori assumptions that are difficult to validate. As a result, different models have different predictive performance during different testing phases across various projects. On the other hand, it is understood that numerous factors may influence systems reliability behavior, including systems complexity, systems development environment, systems development methodology, and so on. Generally, these factors are extremely interactive and exhibit nonlinear patterns. This severely limits traditional modeling methods which are heavily dependent on the assumptions of independence and linearity. Consequently, a single universal model that can provide extremely accurate predictions under all circumstances and without requiring any assumptions is highly desirable.

In recent years, neural-network approach has proven to be a universal approximator for any nonlinear continuous function with an arbitrary accuracy. Consequently, it has become an alternative method in systems reliability modeling, evaluation and prediction. Unlike traditional statistical models, neural network is data-driven, nonparametric weak models and they let "the data speak for themselves". In the existing literature, the use of neural networks is not common in reliability engineering, particularly in engine systems analysis. Liu et al. [1] demonstrated how feed-forward multilayer perceptron (MLP)

*Corresponding author. Tel.: 886 6 5977611; fax: 886 6 5977610.
E-mail address: t530903@ms7.hinet.net.

0951-8320/$ - see front matter
doi:10.1016/j.ress.2005.12.014

networks can successfully identify underlying failure distribution and estimate the parameters. Amjady and Ehsan [2] presented an expert system based on neural networks for evaluating power system reliability. Moreover, Xu et al. [3] applied feed-forward multilayer perceptron neural networks and radial basis function neural networks to forecast engine system reliability. Authors compared the neural-network techniques with the autoregressive integrated moving average (ARIMA) approach. Sensitivity analysis of neural networks was performed and appropriate architectures of neural networks were determined. Notably, Chang et al. [4] applied a hybrid learning neural fuzzy system to forecast engine system reliability. Numerical results demonstrate that the proposed model is able to achieve more accurate forecasting results than ARIMA and generalized regression neural network model (GRNN). Ho et al. [5] conducted a comparative analysis of neural networks and ARIMA techniques in terms of their ability to forecast repairable systems. Experimental results showed that both the recurrent neural networks and multilayer feed-forward neural networks outperform the ARIMA approach in prediction accuracy. Furthermore, Karunanithi et al. [6,7] first proposed using the neural-network approach to predict software reliability. Additionally, Adnan et al. [8,9], Park et al. [10], Aljahdali et al. [11] and Tian and Noore [12] contributed significantly to software reliability prediction using neural networks, and achieved better predictive performance than traditional analytical methods. However, the neural network suffers from a number of weaknesses, which include the need for numerous controlling parameters, difficulty in obtaining a stable solution and the danger of over-fitting.

In 1995, Vapnik [13] developed a neural-network algorithm called support vector machine (SVM), which is a novel learning machine based on statistical learning theory, and which adheres to the principle of structural risk minimization seeking to minimize an upper bound of the generalization error rather than minimize the training error (the principle followed by neural networks). This induction principle is based on the bounding of the generalization error by the sum of the training error and a confidence interval term depending on the Vapnik–Chervonenkis (VC) dimension. Based on this principle, SVM achieves an optimum network structure by striking a right balance between the empirical error and the VC-confidence interval. This balance eventually leads to better generalization performance than other neural-network models [14]. Additionally, the SVM training process is equivalent to solving linearly constrained quadratic programming problems, and the SVM embedded solution meaning is unique, optimal and unlikely to generate local minima. Originally, SVM has been developed to solve pattern recognition problems. However, with the introduction of Vapnik's ε-insensitive loss function, SVM has been extended to solve nonlinear regression estimation problems, such as new techniques known as support vector regression (SVR), which have been shown to exhibit excellent performance [15]. So far there is not any article to use the SVR to construct systems reliability prediction model, so this study intends to employee this new technique and try to construct the systems reliability prediction model.

To construct the SVR model efficiently, SVR's parameters must be set carefully [16,17]. Several past investigations adopted with the method of trail-and-error approaches to tune up the parameters. The usage of these methods must create many SVR models and also take much time to compute the training error. In other words, so far no systemic and architectural methods are available to determine these parameters. Therefore, this study proposes a new approach known as GA-SVR. In GA-SVR, the real-value genetic algorithms (RGA) was employed to determine the optimal parameters of SVR, which were then applied to construct the SVR model. Then, the engine systems reliability was predicted.

The turbocharger is a critical component in a turbocharged diesel engine. Since reliability is one of the most important considerations in diesel engine system design, accurate reliability forecasts can provide a good assessment of engine performance. In most situations, the reliability of manufacturing systems changes over time. These changes can be treated as a time-series process. Instead of evaluating the reliability of complex systems, this study attempts to forecast reliability by analyzing past historical failure data using SVR techniques. This study proposed applying the GA-SVR model to predict engine system reliability owing to its inherent advantages as outlined in Section 2. The proposed approach was compared with the existing neural-network approaches and traditional time series models, such as ARIMA, thus demonstrating that the SVR model is substantially featured with an excellent forecasting capacity.

This study includes seven sections. Section 2 introduces theories related to SVR. Section 3 elaborates on the GA-SVR model proposed in this study. Section 4 describes the data source and experimental settings. Section 5 analyzes the results of RGA and optimizes SVR's parameters. Section 6 discusses and analyzes the experimental results. Section 7 concludes the study and suggests directions for future investigations.

2. Support vector regression

Recently, a regression version of SVM has emerged as an alternative and powerful technique to solve regression problems by introducting an alternative loss function. In the sequel, this version is referred to as support vector regression (SVR). Here a brief description of SVR is given. Detailed descriptions of SVR can be found in Vapnik [13,15,18], Schölkopf and Smola [19] and Cristianini and Shawe-Taylor [20].

The SVR formulation follows the principle of structural risk minimization seeking to minimize an upper bound of