Chapter VIII
Online Approaches to Missing Data Estimation

ABSTRACT

The use of inferential sensors is a common task for online fault detection in various control applications. A problem arises when sensors fail when the system is designed to make a decision based on the data from those sensors. Various techniques to handle missing data are discussed in this chapter. First, a novel algorithm that classifies and regresses in the presence of missing data online is presented. The algorithm was tested for using both classification and regression problems and was compared to an off-line trained method that used auto-associative networks as well as a Hybrid Genetic Algorithm (HGA) method and a Fast Simulated Annealing (FSA) technique. The results showed that the presented methods performed well for online missing data estimation. Second, an online estimation algorithm that uses an ensemble of multi-layer perceptron regressors, HGA and FSA and genetic programming is presented for missing data estimation and compared with a similar procedure that was trained off-line.

INTRODUCTION

Fault detection is one of the most active research areas with several applications. There are a number of challenges faced by online detection and identification systems (Marwala, 2001; Benitez-Perez, Garcia-Nocetti, & Thompson, 2007; Vilakazi & Marwala, 2006; Vilakazi & Marwala, 2007a&b; Marwala, 2007; Hulley & Marwala, 2007). Marwala and Hunt (2000) proposed pseudo modal energies for fault detection in structures. The problem with the method they proposed is that for its successful implementation it had to be continuously retrained and this process required human intervention (Marwala, 2001). One of the biggest problems hindering the performance of online condition monitoring is in dealing with missing data (Abdella & Marwala, 2005). Missing data problems often arise due to sensor failure in online condition monitoring systems. For example, Dhlamini, Nelwamondo, and Marwala (2006) implemented sensor failure compensation techniques for high voltage bushing monitoring using evolutionary computing. This chapter investigates a problem of condition monitoring where fault detection is confounded by missing data. The biggest challenge is that standard neural networks cannot process input
Marwala

data with missing values and hence, cannot perform classification or regression when some input data are missing. More often, there is a limited time between the readings depending on how frequently the sensor is sampled. As a result, missing data problem becomes a huge obstacle in deciding the condition of the machine being monitored from the limited time between readings. For both classification and regression, all decisions concerning how to proceed must be taken during this finite period.

Three general ways have been used to deal with the problem of missing data (Little and Rubin, 1987). The simplest method is known as *listwise deletion*. This method simply deletes instances with missing values (Little & Rubin, 1987). Even though this approach to the problem can be valid for surveys, the major disadvantage of this method is the dramatic loss of information in datasets (Kim & Curry, 1997). Another disadvantage is that it assumes that the observation with missing values is unimportant and can be ignored. For engineering problems where the observations are required for automated decision-making, this option is not viable. The second common technique *imputes* the data by finding estimates of the values and replaces missing entries with these estimates. Various estimates have been used and these estimates include zeros, means and other statistical calculations. These estimates are then used as if they were the observed values. This procedure is valid for engineering problems but the problem is that it results in bad decisions. Another common technique assumes some model for the prediction of the missing values and uses the maximum likelihood approach to estimate the missing values (Little & Rubin, 1987; Nelwamondo, Mohamed, & Marwala, 2007).

Much research has been done to find new ways of approximating missing values. Among others, Abdella and Marwala (2006) as well as Mohamed and Marwala (2005) used neural networks together with Genetic Algorithms (GA) to approximate missing data. Qiao, Gao, and Harley (2005) used neural networks and Particle Swarm Optimization (PSO) to keep track of the dynamics of a power plant in the presence of missing data. In the aforementioned examples, auto-associative neural networks were used together with GA or PSO to predict the missing values and to optimize the prediction to be as accurate as possible. On the other hand, Yu and Kobayashi (2003) used semi-hidden Markov models to predict missing data in mobility tracking whereas Huang and Zhu (2002) used a pseudo-nearest-neighbor approach for missing data recovery of Gaussian random data sets. Nauck and Kruse (1999) and Gabrys (2002) have also used neuro fuzzy techniques in the presence of missing data. A different approach was taken by Wang (2005) who replaced incomplete patterns with fuzzy patterns. Along with fuzzy patterns, the patterns without missing values were used to train the neural network. In Wang’s model, the neural network learns to classify without actually predicting the missing data.

This chapter introduces techniques to handle missing data in an online system. Online systems require continuous learning. Many techniques have been proposed to deal with this problem. Lunga and Marwala (2006) proposed an ensemble of classifiers for online classification of a time series. Mohamed, Marwala, and Rubin (2007) applied online learning for adaptive protein classification. An approach where no attempt was made to recover the missing values was also presented. In this approach, both classification and regression techniques were considered. For classification, an approach that uses an ensemble of classifiers, also called a *committee* as in Chapter VII, to classify even in the presence of missing data was presented. Unlike Wang (2005), it did not attempt to replace missing patterns with anything. Instead, the network was trained with various subsets of available data. The algorithm was further extended to a regression application where Multi-Layer Perceptrons (MLP) were used in an attempt to get the correct output with limited input variables.

The last part of this chapter extends the technique used by Dhlamini, Nelwamondo, and Marwala (2006) for the classification of faults in transformer bushings in the presence of missing data. In Chap-
20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage: 
www.igi-global.com/chapter/online-approaches-missing-data-estimation/6800?camid=4v1

www.igi-global.com/e-resources/library-recommendation/?id=1

Related Content

The Formal Design Model of an Automatic Teller Machine (ATM)
Yingxu Wang, Yanan Zhang, Philip C.Y. Sheu, Xuhui Li and Hong Guo (2010). International Journal of Software Science and Computational Intelligence (pp. 102-131).
www.igi-global.com/article/formal-design-model-automatic-teller/39108?camid=4v1a

Effects of a Preventive Warning Light System for Near-Miss Incidents
www.igi-global.com/article/effects-of-a-preventive-warning-light-system-for-near-miss-incidents/199017?camid=4v1a

Improving the Efficiency of Large-Scale Agent-Based Models Using Compression Techniques
www.igi-global.com/chapter/improving-efficiency-large-scale-agent/67299?camid=4v1a

Hierarchical Function Approximation with a Neural Network Model
Luis F. de Mingo, Nuria Gómez, Fernando Arroyo and Juan Castellanos (2009). International Journal of Software Science and Computational Intelligence (pp. 67-80).
www.igi-global.com/article/hierarchical-function-approximation-neural-network/34089?camid=4v1a