International Journal of Hybrid Intelligent Systems 10 (2013) 191–203
DOI 10.3233/HIS-130176
IOS Press

An automatic method for construction of ensembles to time series prediction

Tiago P.F. Lima* and Teresa B. Ludermir
Centro de Informática, Universidade Federal de Pernambuco, Recife, Brazil

Abstract. We present here a work that applies an automatic construction of ensembles based on the Clustering and Selection (CS) algorithm for time series forecasting. The automatic method, called CSELM, initially finds an optimum number of clusters for training data set and subsequently designates an Extreme Learning Machine (ELM) for each cluster found. For model evaluation, the testing data set are submitted to clustering technique and the nearest cluster to data input will give a supervised response through its associated ELM. Self-organizing maps were used in the clustering phase. Adaptive differential evolution was used to optimize the parameters and performance of the different techniques used in the clustering and prediction phases. The results obtained with the CSELM method are compared with results obtained by other methods in the literature. Five well-known time series were used to validate CSELM.

Keywords: Clustering and selection, ensembles, extreme learning machine, self-organizing maps, time series forecasting, adaptive differential evolution

1. Introduction

Time series forecasting has been used in several real world problems to support the decision-making process [6]. Various disciplines rely heavily on the ability to forecast. For example, time series analysis is used extensively in economics and finance to predict inflation levels, stock market prices and currency exchange rates. Traditional time series approaches tend to focus on methods first proposed by Box and Jenkins [6] but several models can be used to forecast a time series. However, there is no single model considered the best [3]. Selecting the most adequate model for a given series may be a difficult task, depending on the candidate models and on the characteristics of the series. An approach that can be used for model selection is the development of expert systems [9]. The knowledge used in these systems is extracted from forecasting experts and it is expressed in the form of rules. In this context, we highlight the Rule-Based Forecasting system [9], which implemented an expert system with rules, associating time series features to the available models. Although these systems can express knowledge in a practical and reusable way, the process of knowledge acquisition depends on human experts, which are often expensive. In this scenario, machine learning techniques are an interesting alternative to acquire knowledge. These techniques can be used to automatically learn from data, leading to potential improvement of performance, easy adaptability to new types of data, and a reduced need for experts [3].

Artificial Neural Networks (ANNs) are examples of effective learning machines in time series forecasting [37]. Nevertheless, many real-world problems are too large and complex for a single ANN to solve alone. In an attempt to improve performance, a common approach is to combine multiple ANNs, forming ensembles [15]. There are several theoretical and practical reasons for ensembles [25]. First, ensembles or committees, exploit the idea that a pool of different learning machines, also referred as to experts, can offer complementary information about the patterns to be predicted, improving the effectiveness of the overall pro-

*Corresponding author: Tiago P.F. Lima, Centro de Informática, Universidade Federal de Pernambuco, Av. Jornalista Anibal Fernandes, s/n, Cidade Universitária, 50740-560, Recife-PE, Brazil. Tel.: +55 81 2126 8430; E-mail: tpfl2@cin.ufpe.br.