Chapter 3.13
Electricity Load Forecasting Using Machine Learning Techniques

Manuel Martín-Merino Acera
University Pontificia of Salamanca, Spain

ABSTRACT

Electricity load forecasting has become increasingly important due to the strong impact on the operational efficiency of the power system. However, the accurate load prediction remains a challenging task due to several issues such as the nonlinear character of the time series or the seasonal patterns it exhibits. A large variety of techniques have been proposed to this aim, such as statistical models, fuzzy systems or artificial neural networks. The Support Vector Machines (SVM) have been widely applied to the electricity load forecasting with remarkable results. In this chapter, the authors study the performance of the classical SVM in the problem of electricity load forecasting. Next, an algorithm is developed that takes advantage of the local character of the time series. The method proposed first splits the time series into homogeneous regions using the Self Organizing Maps (SOM) and next trains a Support Vector Machine (SVM) locally in each region. The methods presented have been applied to the prediction of the maximum daily electricity demand. The properties of the time series are analyzed in depth. All the models are compared rigorously through several objective functions. The experimental results show that the local model proposed outperforms several statistical and machine learning forecasting techniques.

DOI: 10.4018/978-1-60960-818-7.ch3.13

INTRODUCTION

Electricity load forecasting has become increasingly important in the last years for the industry due to the deregulation of the electricity markets. The supply industry requires forecasts with lead times that range from short term (hours or days ahead) to long term (years ahead). The accurate short term load forecasting has a significant impact on the operational efficiency of the power system (Hippert, 2001; Khotanzad, 1997). Most utilities believe that improving the accuracy of the electricity load forecasting allows to improve the generation scheduling and the power purchase/ sales decisions (Hobbs, 1998). A better scheduling will result in economic savings avoiding the unnecessary start-up costs for generators that are not actually needed, or avoiding costly combustion turbine generation or purchases of spot power when too little capacity has been committed. Avoiding unneeded power purchases and the identification of favorable sales opportunities will be profitable for the suppliers.

However, the accurate load prediction remains a difficult task. First, the time series is strongly non-linear and exhibits several levels of seasonality and periodicity (Hippert, 2001; Chen, 2004). Secondly, there are many exogeneous variables that must be considered such as the temperature, wind speed or the humidity. Moreover, the relation among the exogeneous variables and the load demand is frequently difficult to model.

A large variety of techniques have been proposed to this aim such as statistical models (Taylor, 2006), fuzzy systems (Papadakis, 1998) or Artificial Neural Networks (ANNs) (Hippert, 2001). Althought, empirical studies (Hobbs, 1998) suggest that the Artificial Neural Networks are able to reduce significantly the errors in the electricity load forecasting, they are prone to overfitting (Hippert, 2001). More recently, several authors have applied the Support Vector Machines (SVM) to the electricity load forecasting (Schölkopf, 1999; Lau, 2008) with remarkable results. SVM are powerful non-linear techniques proposed under a soundness statistical theory that keep a higher generalization ability than most ANNs (Vapnik, 1998). Unfortunately, the properties of the load time series change locally with time due to seasonal effects, holidays and other factors (Chen, 2004; Marín, 2001). For instance, the load patterns differ significantly in winter and summer seasons or in weekends and working days. Therefore global models such as the SVM are not suitable to predict accurately the load demand.

It has been suggested in the literature that splitting the time series into homogeneous regions helps to improve the forecasting accuracy (Chen, 2004; Marín, 2001; Lamedica, 1996). Thus (Marín, 2001; Lamedica, 1996) have applied Artificial Neural Networks such as the Self Organizing Maps (SOM) to identify days of similar load profiles. However, human experts are needed to classify the SOM prototypes which is a serious drawback. In (Dablemont, 2003) a SOM is employed to create Voronoi regions for input and output spaces. Next a frequency table is built that relates both spaces. However the results are not satisfactory when the sample size is not large. Besides, the neighborhood relations induced by the SOM are not considered. Finally, the method proposed by (Cherkassky, 1996) takes advantage of the neighborhood relations induced by the SOM to adjust a locally weighted linear regression. However the SOM is organized without considering the target to be predicted. Therefore relevant information is lost.

In this chapter, we analyze the performance of the Support Vector Machines (SVM) in the problem of electricity load forecasting. Next, we propose a new method that integrates the SVM and the Self Organizing Maps (SOM) in order to exploit the local character of the time series. To this aim, the time series is first splited into homogeneous regions using the SOM. Next, an SVM is adjusted locally in each homogeneous regions. The partition of the input space considers the target to be predicted, avoids the overfitting and takes advantage of the neighborhood relations induced by the SOM. Finally, the performance of the algorithms is illustrated in the challenging problem of maximum daily electricity demand.