Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 122 (2017) 269–275

Procedia Computer Science

www.elsevier.com/locate/procedia

Information Technology and Quantitative Management (ITQM 2017)

Power Demand Forecasting and Application based on SVR

Qing Guo [a], Yuyao Feng [b,c], Xiaolei Sun [b,c,*] , Lijun Zhang [b]

[a] *State Grid Energy Research Institute, Beijing 102209, China*
[b] *Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100190, China*
[c] *School of Public Policy and Management, University of Chinese Academy of Sciences, Beijing 100049, China*

Abstract

Power as one of the most important energy to promote the development of national economy, plays an essential role in the normal operation of all aspects of society. Because of its production and use are difficult to store in large quantities, it is necessary to forecast future demand, which will become an important basis for making power development plans. Considering the complex non-linear relationship between power demand and its influencing factors, it is difficult to describe it accurately with the traditional mathematical models. In this paper, we select six major influencing factors and use the support vector machine to predict future power demand. The prediction accuracy is improved by parameter optimizing. At the same time, the simulation experiment of Shandong Province is conducted to further verify the validity and feasibility of the model.

Peer-review under responsibility of the scientific committee of the 5th International Conference on Information Technology and Quantitative Management, ITQM 2017.

Keywords: Power system; Demand forecasting; Support vector regression; Parameter optimization

1. Introduction

Accurate power demand prediction is a key step in the study of power system. Given that power cannot be stored in an economical and viable way, there must be a balance between supply and demand [1], in which power demand forecasting plays a crucial role. Owing to the importance of power demand forecasting, a wide variety of models have been proposed in the past. Among them, the following two categories are the most common: one is traditional forecasting technology, such as regression analysis, time series analysis based on ARIMA [2,3]. The other is modeling the problem through nonlinear intelligent prediction methods, mainly including neural network, gray system, fuzzy forecasting and so on [4-8].

In recent years, statistical learning theory has been widely concerned and has become a popular research object in the field of machine learning. The support vector machine (SVM), which is derived from this theory has made

* Corresponding author. Tel.: +86-010-59358806; fax: +86-010-59358608
E-mail address: xlsun@casipm.ac.cn

1877-0509
Peer-review under responsibility of the scientific committee of the 5th International Conference on Information Technology and Quantitative Management, ITQM 2017.
10.1016/j.procs.2017.11.369

great contributions to machine learning and other application fields. Although the SVM approach has been applied to power demand [9,10], there is still a long way to go to achieve the best predicting outcomes. Data sources, data pre-processing and parameter optimization are likely to be the factors that impede its development. So, this paper presents a more rigorous data preprocessing method and a more effective parameter optimization method, which realizes the improvement of SVM method in the field of power demand forecasting.

The rest of this paper is organized as follows: In section 2, models used in this paper are briefly introduced. In section 3, empirical analysis is operated using statistic data of Shandong Province. At last, conclusions are given in section 4.

2. Models

SVM was first proposed by Corinna Cortes and Vapnik in 1995. It is a kind of machine learning method based on statistical learning theory, mainly used in classification and regression analysis [11]. In simple terms, the basic idea behind SVM is raising the dimension and linearization. It has many unique advantages in solving small sample, nonlinear and high dimensional pattern recognition.

If the predicted variable is discrete, it is called classification, and if the predicted variable is continuous, it is called regression. In this paper, support vector regression(SVR) will be used to deal with experiment data. In this section, we will briefly describe the mathematical derivation of SVR.

2.1. Linear case

For the linear case, SVM first considers using the linear function $f(x)=(\omega\cdot x)+b$ to fit (x_i,y_i) $i=1,2,\ldots l$, where $x_i\in R^n$ is the input variable, $y_i\in R^n$ is the output variable. Then, our goal is to find the optimal solution for w and b, so that the function $f(x)$ can approximate future values accurately. SVM aims to construct a hyperplane with maximum the edge distance between support vectors. By calculation, the edge distance can be expressed as $2/\|\omega\|$, formally this can be converted to a convex optimization problem by minimizing $\frac{1}{2}\|\omega\|^2$ [12].

When it comes to regression, the above convex optimization problem is feasible. But in reality, the experimental data will be affected by noise, resulting in errors [9]. So, in order to measure the experimental error. SVR approximates all (x_i,y_i) pairs with ε precision, slack variable ξ_i,ξ_i^* and the penalty coefficient C also be introduced to constraint the experimental error. After the constraint conditions have been added, the objective function can be expressed as follows:

$$\min \quad \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{l}(\xi_i+\xi_i^*)$$

$$s.t. \quad \begin{Bmatrix} y_i-(\omega\cdot x)-b\le\varepsilon+\xi_i \\ (\omega\cdot x)+b-y_i\le\varepsilon+\xi_i^* \\ \xi_i,\xi_i^*\ge 0 \end{Bmatrix}$$

The next step, we use Lagrange function to normalize the objective function, the formula is defined as:

$$L=\frac{1}{2}\|\omega\|^2+C\sum_{i=1}^{l}(\xi_i+\xi_i^*)-\sum_{i=1}^{l}\alpha_i(\varepsilon+\xi_i-y_i+(\omega\cdot x)+b)-$$
$$\sum_{i=1}^{l}\alpha_i^*(\varepsilon+\xi_i+y_i-(\omega\cdot x)-b)-\sum_{i=1}^{l}(\eta_i\xi_i+\eta_i^*\xi_i^*)$$

Among them, $\alpha_i,\alpha_i^*,\eta_i,\eta_i^*\ge 0$ are Lagrange multipliers.

After the dual transformation, the process of minimizing the objective function may be equivalent to the equation $max_{\alpha,\eta}\min_{w,b,\xi}$ L. After calculating the partial derivatives of each parameter, the objective function