VOL.2 NO.2 September 2013 - ISSN 2233 – 1859

UOIuBIH
ORSinBIH
Operations Research Society in Bosnia and Herzegovina

Southeast Europe Journal of Soft Computing

Available online: www.scjournal.com.ba

Comparison of Machine Learning Methods for Electricity Demand Forecasting in Bosnia and Herzegovina

Galib Sikirić, Samir Avdaković, Abdulhamit Subasi

International Burch University, Faculty of Engineering and Information Technologies, Francuske Revolucije bb. Ilidza, Sarajevo, 71000, Bosnia and Herzegovina

Article Info

Article history:
Recivied 17 Sep.2013
Recivied in revised form 17 Oct 2013

Keywords:
Short Term Load Forecasting, Active Power, Reactive Power, Apparent Power, Machine Learning Methods.

Abstract

Electricity demand forecasting is one of the most important components in the power system analysis. Furthermore, it is difficult and complicated process to forecast energy consumption. This study deals with modeling of the electrical energy consumption in Bosnia and Herzegovina in order to forecast future consumption of electrical loads based on temperature variables using machine learning methods. We used three different machine learning methods for analyzing short term forecasting. The methods were trained using historical load data, collected from JP Elektroprivreda electrical power utility in BiH, and also considering weather data which is known to have a big impact on the use of electric power. Comparing the results it was seen that prediction for 500 hours is pretty good in range from 92,92% for reactive power till 98.84% for active power. Four different parameters were analyzed mean absolute error, root mean squared error, relative absolute error and root relative square error. The best results for apparent power were gotten with linear regression and are presented as for mean absolute error 9.84, root mean squared error 13.62, relative absolute error 14.06%, root relative squared error 14.39%. It is also seen from the results that, the short term power consumption can be predicted which is important for maintaining of the voltage at the consumer side.

1. INTRODUCTION

Electricity demand forecasting is one of the most important components in the power system analysis and development and present a very difficult and complicated process. In the context time horizon of the planning or forecasting, the short-term, mid-term and long-term planning or forecasting is a usual classification that can be found in the literature. Short-term planning is usually an hourly, daily or weekly operating planning of power system, and one of the most important aspects during this process is a quality and accurate forecasting of electricity demand. Accurate short-term forecasts are very important in the electricity market because it could have an affect the price of electricity that will be supplied to customers. On the other hand, mid-term power system operating planning is usually prepared on a time horizon of one or more months, while long-term planning are usually prepared on annual basis, and long-term forecasting of electricity demand is usually presented for several years and it is a input into the process of building a new generating units planning. Today, a vast number studies and papers provide a different electricity demand forecasting methodology, which confirm a great interest professional and scientific community in this field. During forecasting, engineers trying to identify the main factors affecting the consumption of electricity, such as growth of GDP, demographic and climate change, standards and habits of population, etc. One of the most important factor in the short-term and mid-term forecasting and requires special attention is the air temperature. The different approaches to determining the influence of air temperature on electricity consumption are presented in Refs. [1-5] and usually, all analyzed cases confirm the high interrelationships between these variables. Also, during the process of forecasting, the ambient temperature, human social activity and other variables provides important information that forecast results can make a lot more accurate.

In this paper three different models of Machine learning methods for analyzing short term load forecasting are proposed. This paper is organized as follows. Section 2 briefly outlines the methods used in this paper, while Section 3 provides data used during process of forecasting. Results and discussions are presented in Section 4 and finally, Section V presents the conclusions of this paper.

2. MACHINE LEARNING TOOLS

For machine learning tool we can simply say that it is a branch of artificial intelligence that can learn from data construct and study different types of systems. A number of machine learning models exist like artificial neural networks (ANN), support vector machine (SVM), decision tree learning, linear regression etc. Learning with machine tools can be unsupervised, semi-supervised, reinforcement learning, learning to learn and supervised learning where the last is one of the most popular and has been extremely explored. The task of machine tools is to train a machine to identify the connection between the examples and the desired outputs (class labels), while an external teacher (supervisor) provides the set of examples related with desired outputs. After the machine has been trained, it becomes ready to provide the output for a new unseen example. In our work we used SMO (Sequential Minimal Optimization) which is an improved training algorithm for SVMs. As it is shown in [7] SMO type of SVM represents a set of simple algorithms that can break down large QP (quadratic programming) into a series of smaller QP problems. The advantage of SMO type of SVM is that it can solve the Lagrange multipliers analytically and it represent supervised learning algorithm used for classification and it is a fast implementation of SVM. SMO type of SVM constructs a hyperplane or a set of hyperplanes in an n-dimensional space which can be used for classification. For the separation we can say it is good if the hyperplane has the largest distance to the nearest training data points of any class and we have in general the larger the margin the lower generalization error of classifier. Two different kernels: Polynomial kernel and the PUK kernel were employed in this study. One of the machine learning tools that was also used in this paper is linear regression. It is the oldest and most broadly used predictive technique in the area of the machine learning. In prediction and forecasting if we have an experiential data set of y and X linear regression can be used to built-in a predictive model. After we developed a model, and if some value of X is given without its additional value of y, the built-in model can be used to make a prediction of the value of y.

3. THE DATA

The input data that we are dealing with in this paper is composed of active, reactive, apparent power and temperature. It was collected from a part of the Bosnia and Herzegovina (B&H) electric power system and provides supplying about 700,000 customers. Basic characteristics of this consumption are relatively low electricity consumption due to a very low population standards and poor economic development, where households have dominant impact in total electricity consumption. Temperature fluctuations of this area are in quite a vast range and for the period of one year they are in the range from -12 up to 30 ° C. Hourly values of electricity demand and temperature are presented in Fig.1 and come from the EPC Elektroprivreda B&H-Sarajevo for electricity demand and the Federal Hydro meteorological Institute of B&H for temperatures. It is apparent close relationship between temperature and electricity consumption.

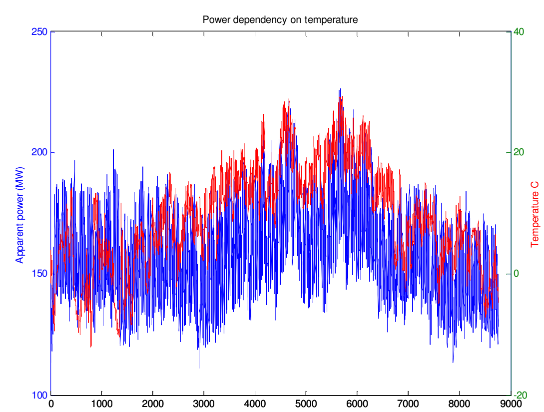

Figure 1. Hourly values of electricity demand and temperature in B&H

4. RESULTS AND DISCUSSION

This research study was perform on the previously mentioned data were 2500 instances (hours) are taken for training and 500 instances (hours) for predicting. In the Table 1, we can see the results of the experiment that was done in weka for active power. As we can see we have very high accuracy where the best solution is given for linear regression 98.84%. As it was expected since from the input data given in Fig.1 temperature data is following apparent power consumption. If we look at the errors basically the best prediction is linear regression for active power while for the reactive power the smallest errors we have for the SMO type of SVM (PUK kernel) method. The reason is simple since the consumers are mainly houses the active energy is mostly consumed using coolers and heaters as temperature decreases and rises up.

Table 1. Results for the active power

Coefficients/ methods	Linear Regression	SVM (PolyKernel)	SVM (PUK)
Correlation coefficient	0.9884	0.9883	0.9867
Mean absolute error	9.9228	10.282	11.1817
Root mean squared error	13.7546	14.2138	14.5888
Relative absolute error	13.7908%	14.29%	15.5404%
Root relative squared error	13.7548%	14.214%	14.5889%
Total Number of Instances	500	500	500

Table 2. Results for reactive power

Coefficients/ methods	Linear Regression	SVM (PolyKernel)	SVM (PUK)
Correlation coefficient	0.9292	0.9326	0.9468
Mean absolute error	4.2766	4.1881	3.8546
Root mean squared error	5.593	5.5144	4.9273
Relative absolute error	31.365%	30.7162%	28.2701%
Root relative squared error	33.8226%	33.3471%	29.7968%
Total Number of Instances	500	500	500

5. CONCLUSIONS

In this paper we tried to predict active and reactive power for 500 hours (20 days) and we succeeded with a very high accuracy. Better prediction is seen for active power since as we said the consumers are mainly active but predicting the reactive power is problem that needs to be more analyzed in the further research. These predictions are not that important since they are too short but still they can help us in daily consumption when we want to predict the peaks during the day. Further research will be focused on long term prediction where we need to include more parameters such as economic and demographic parameters. And these long term predictions are more valuable since they can help us in predicting whether we need to build more power plants in certain areas or not.

REFERENCES

1. Henley A, Peirson J. Non-Linearities in electricity demand and temperature: Parametric Versus Non-Parametric Methods. Oxford Bulletin of Economics and Statistics 1997; 59(1): 149-162.

2. Lee CC, Chiu YB. Electricity demand elasticities and temperature: Evidence from panel smooth transition regression with instrumental variable approach. Energy Economics 2011; 33(5): 896-902.

3. Kervinen T. Estimation of the temperature dependency of the Finnish electricity consumption. Helsinki University of Technology- Independent Research Project in Applied Mathematics 2008; (see also www.sal.tkk.fi/publications/pdf-files/eker08.pdf)

4. Pilli-Sihvola K, Aatola P, Ollikainen M, Tuomenvirta H. Climate change and electricity consumption-witnessing increasing or decreasing use and costs?. Energy Policy 2010; 38(5): 2409-2419.

5. Gupta E. Climate change and the demand for electricity: A non-linear time varying approach. Indian Statistical Institute-Delhi 2011; (see also www.isid.ac.in/~pu/.../dec.../EshitaGupta.pdf)

6. Alfares H.K, Nazeeruddin M. Electric load forecasting: literature survey and classification of methods. Int Jour of Systems Science 2002; 33(1): 23-34.

7. John C. Platt, Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machine, Microsoft Research, Technical Report MSR-TR-98-14, April 21, 1998.