SERBIAN JOURNAL OF ELECTRICAL ENGINEERING
Vol. 8, No. 1, February 2011, 73-83

UDK: 007.52

Application of SVM Methods for Mid-Term Load Forecasting*

Miloš Božić[1], Miloš Stojanović[2]

Abstract: This paper presents an approach for the medium-term load forecasting using Support Vector Machines (SVMs). The proposed SVM model was employed to predict the maximum daily load demand for the period of a month. Analyses of available data were performed and the most important features for the construction of SVM model are selected. It was shown that the size and the structure of the training set may significantly affect the accuracy of predictions. The presented model was tested by applying it on real-life load data obtained from distribution company "ED Jugoistok" for the territory of city Niš and its surroundings. Experimental results show that the proposed approach gives acceptable results for the entire period of prediction, which are in range with other solutions in this area.

Keywords: Load forecasting, Support vector machines, Regression.

1 Introduction

Load forecasting has always been an important issue for economic and reliable operation of electric utilities. Load forecasting can be related to the load of the whole country or on load related to a dealer or customer group, for the period of the next hour, day, month or all year. Depending on the length of the prediction period, it can be divided into three types: long, medium and short term. Long-term load forecasting is related for the period from 5 to 20 years and is particularly important in terms of deciding when to conduct the extension of the electricity network. Mid-term load forecasting is used to estimate the load in the winter or summer period, and usually refers to a period from several days to several weeks, or several months. Short-term load forecasting mainly covers the period of one week, and refers to the assessment of load per hour during the day.

In recent years much effort has been employed to explore possibilities of applying machine learning techniques to the problem of forecasting electricity

[1]University of Niš, Faculty of Electrical Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia;
E-mail: milos1bozic@gmail.com
[2]Technical College of vocational studies, Aleksandra Medvedeva 20, 18000 Niš, Serbia;
E-mail: milos.stojanovic@vtsnis.edu.rs
*Award for the best paper presented in Section *Power Engineering*, at Conference ETRAN 2010, June 7-11, Donji Milanovac, Serbia.

demand. The most commonly used machine learning techniques for the electricity demand forecast are SVMs [1] and Artificial Neural Networks (ANNs) [2].

In this paper, an approach for prediction of maximum daily load demands for a period of a month, using SVMs is presented. For the training and testing of proposed model, load demands for the territory of city Niš for the period of 2008 and 2009 are used, as well as temperatures and wind speed in the same period. The paper is organized as follows. In Section 2, SVM techniques are presented. In Section 3 a brief analysis of available data is made, in order to identify features for the formation of SVM model. In Section 4 experimental results are presented and discussed. Section 5 contains the conclusion and proposes a direction for further research.

2 SVMs

SVMs are proposed by Vapnik in [3] in order to resolve the issue of data classification. Two years later, the version of SVMs is proposed that can be successfully applied to the non-linear regression problems. This method is called Support Vector Regression (SVR) and it is the most common form of SVMs. SVMs are based on the principle of structural risk minimization (SRM), which is proved to be more efficient than the empirical risk minimization (ERM), which is used in neural networks. SRM minimizes an upper bound of expected risk as opposed to ERM that minimizes the error on the training data [4].

The goal of SVR is to generate a model which will predict unknown outputs based on the known inputs. In the training phase the formation of the model is performed based on the known training data (x_1, y_1), (x_2, y_2), …, (x_n, y_n), where x_i are input vectors, and y_i output scalars associated to them. In the test phase the trained model on the basis of new inputs $x_1, x_2, ..., x_n$ makes prediction of unknown outputs $y_1, y_2, ..., y_n$.

Consider a known training set $\{x_k, y_k\}$, $k = 1,\dots,N$ with input vectors $x_k \in R^n$ and output scalars $y_k \in R$. The following regression model can be built by using non-linear mapping function $\varphi(\cdot): R^n \to R^{n_h}$ which maps the input space into a high-dimensional feature space and constructs linear regression in it, expressed in (1):

$$y(x) = \omega^T \varphi(x) + b\,, \tag{1}$$

where ω represents the weight vector and b is a bias term. Optimization problem is formulated in primal space in (2):

$$\min_{\omega, b, \xi, \xi^*} \frac{1}{2}\omega^T \omega + C\sum_{i=1}^{N}(\xi_i + \xi_i^*)\,, \tag{2}$$

subject to constrains expressed in (3):

$$\begin{aligned} & y_i - (\omega^T \varphi(x_i) + b) \le \varepsilon + \xi_i, \\ & (\omega^T \varphi(x_i) + b) - y_i \le \varepsilon + \varepsilon_i^*, \end{aligned} \quad \xi_i, \xi_i^* \ge 0, i = 1, 2, ..., N, \tag{3}$$

where x_i is mapped in the high-dimensional vector space with mapping φ, ξ_i is the upper limit of training error and ξ_i^* the lower. C is a constant which determines the "cost of error", i.e. determines the tradeoff between the model complexity and the degree to which deviations larger than ε are tolerated. Parameter ε controls the width of ε insensitive zone, used to fit the training data [5]. Parameters C and ε are not known in advance and must be determined by some learning algorithm applied on the training set (e.g. Grid – Search and Cross-Validation). The goal of SVR is to place as many input vectors x_i inside the tube $\left| y - (\omega^T \varphi(x) + b) \right| \le \varepsilon$, which is shown in Fig. 1. If x_i is not inside the tube, an error occurs ξ_i or ξ_i^*.

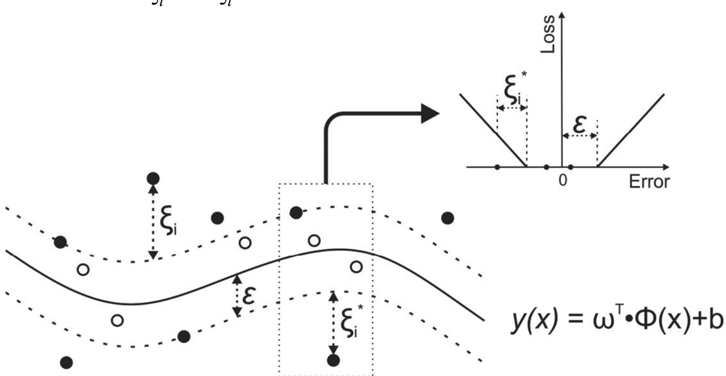

Fig. 1 – ε *tube of nonlinear* SVR.

In order to solve optimization problem defined with (2) and (3), it is necessary to construct dual problem using Lagrange function and calculate weights ω and bias b. The resulting SVM model for regression in dual form is represented in (4), where α_i and α_i^* are Lagrange multipliers and $K(x_i, x)$ represents kernel function, defined as dot product between $\varphi(x_i)^T$ and $\varphi(x)$.

$$y(x) = \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) K(x_i, x) + b\,. \tag{4}$$

Kernel functions enable computation of the dot product in a high-dimensional feature space by using data inputs from original space, without explicitly computing $\varphi(x)$. Commonly used kernel function in non-linear regression problems, one that is employed in this study, is radial basis function (RBF) represented in (5):

$$K(x_i, x) = \varphi(x_i)^T \varphi(x) = e^{-\gamma\|x_i - x\|^2}, \quad \gamma > 0, \tag{5}$$

where γ represents kernel parameter, which should be also determined by learning algorithm. More about SVR can be found in [6, 7].

3 Data Analysis and Feature Selection

Data analysis:

Avaliable data for construction of forecasting model consist of following features:

- electrical load (active power) measured on each 15 minutes,
- average daily temperature,
- average daily wind speed,
- average daily humidity,
- average daily value of atmospheric pressure,
- dates of holidays and local events,

for the territory of city Niš for 2008 and 2009.

Before feature selection for the model, analyses of available data were performed in order to identify features that have the greatest impact on load behavior. On that occasion certain relations between electric load and other parameters are observed, such as weather conditions or local holidays.

Properties of load demand:

In Fig. 2 maximum daily loads from 2008 to 2009 are presented. Relations between electricity usage and weather conditions in different seasons are observed from data on Fig. 2. Load demand in winter period is higher compared to summer period. Additionally, it can be observed that the seasonal load pattern is almost the same for both years.

Also, load pattern periodicity exists in every week. Usage of electricity is relatively constant during week days and drops in the weekend, Fig. 3. That is not the case when the holidays coincide with the weekends or when significant temperature differences occur in the weekend.

Weather influence:

Relations between weather and load demand was shown in previous works on short term load forecasting [1, 2]. Weather conditions influence on analysis of load demand pattern may include temperature, wind speed, humidity, pressure and illumination, etc. In this paper two weather parameters are considered, temperature and wind speed.

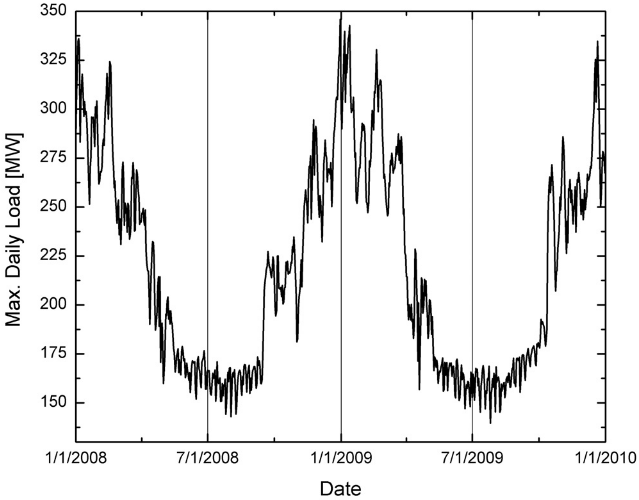

Fig. 2 – *Load pattern during period of two years.*

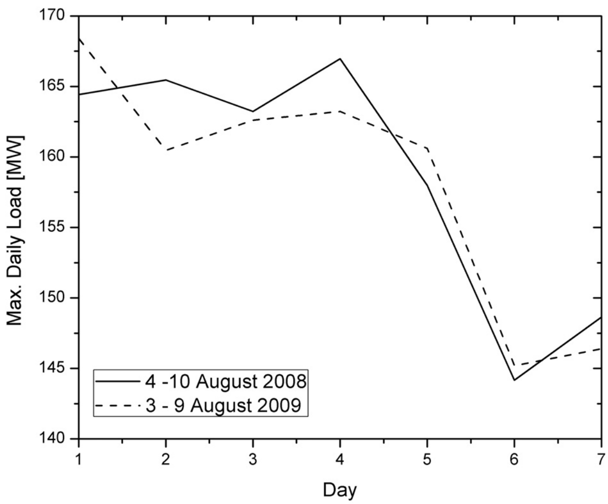

Fig. 3 – *Load pattern during the week.*

Holiday influence:

Holidays and other local events also may affect the load demand. These events are usually local and their influences highly depend on the customs of the area. Major holidays such as Christmas or New Year have more influences on load demand then other holidays.

Feature selection:

In this section a brief analysis is given which features should be included in the data set for the formation of SVM model.

Calendar features. We pointed out earlier that the load demand on week days is relatively constant and higher than in the weekend. Also, the load demand on holidays is lower than on non-holidays. This information (weekdays and holidays) is useful for feature in training entries. Also, many authors [1, 2, 8] have used the calendar information to model the problem.

Temperature feature. The temperature is another useful feature, since load demand and temperature have causal relation in between. Weather information which includes temperature, wind speed, sky cover, humidity and etc., has also been used in most load forecasting methods. But there is one difficulty: for mid-term load forecasting, temperature data for several weeks away are needed.

Wind speed feature. Beside temperature, wind speed is another weather parameter which has important influence on load demand. As for temperature, if we want to encode the wind speed in our training entries, we will also need forecasted wind speed data for several weeks away.

Time series features. Another information which is taken into consideration as the feature is the past load demand. With this approach, concept of time-series [8] is introduced into model. If y_i is the target value for prediction, the vector x_i includes several previous target values $y_{i-1},\ldots,y_{i-\Delta}$ as features. In the training phase all y_i are known but for future prediction, $y_{i-1},\ldots,y_{i-\Delta}$ can be values from previous predictions. This approach is used because, the past load demand could affect and imply the future load demand. In this paper, $\Delta = 7$ is used, which means that seven past daily load demands are used for features.

Model formation:

The various factors that affect the load demand are analyzed and appropriate features are chosen for the SVM forecasting model. Proposed model is shown in Fig. 4. Input vectors are composed of the following features:

- Maximum daily loads for past seven days (P_{i-k}), $k = 1,\ldots,7$;
- Average daily temperatures (T_i);
- Average daily wind speed (W_i);
- Day of the week (DOW).

In addition to the selected features, in order to get "well trained" SVM model, the following parameters should be determined:

- Kernel function and its parameters,
- "cost of error" – C,
- width of tube – ε .

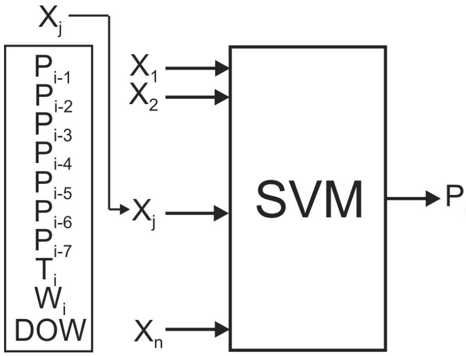

Fig. 4 – *Proposed architecture.*

Radial Basis Function (RBF) is selected for kernel function, defined in (5).

A simple method to determine SVM parameters is Grid – Search [6] over all combinations of parameters. Instead of evaluating every possible parameter combination, which would be time consuming, a grid using equidistant steps limits the search complexity. In our experiments couples C and γ are searched exponentially in range $C = 2^{-5} - 2^{20}$ and $\gamma = 2^{-15} - 2^{3}$ with steps 0.5 and 0.1.

For evaluation how model performs prediction with different pairs of parameters C and γ, k – fold Cross – Validation [6] procedure is used. The training set is randomly divided into training and testing parts, in relation 1: k. Then the learning algorithm is applied to the training part with current values of C and γ. Then, evaluation of the prediction quality is performed on the testing part. This procedure is repeated k times and a pair C and γ is selected, with which the best predictions are achieved.

In order to speed up model formation ε was not included in Grid – Search procedure, so ε was tuned “manually”. After selecting features and determining SVM parameters, forecasting model is constructed, according to (4).

4 Experimental Results

In electricity load forecasting, the prediction accuracy is evaluated using Mean Absolute Percentage Error (MAPE) [10]. This error is defined with (6):

$$MAPE = 100 \cdot \frac{1}{n} \sum_{i=1}^{n} \left| \frac{P_i - \hat{P}_i}{P_i} \right|, \tag{6}$$

where P_i and $\hat{P}_i$ are the real and the predicted value of the maximum daily electrical load on the i^{th} day, and n is the number of days in the month.

The model evaluation was made based on the loads for the month of January 2010, while for the model training different subsets of data are selected from the basic training set. For the model training, the training sets are formed of the following data:

- data for 2009,
- data for winter 2009,
- data for December, January and February 2009,
- data for summer 2009.

Table 1
MAPE values for different training segments.

Training set	2009	Winter 09	Dec-Jan-Feb	Summer 09
MAPE [%]	4.25	4.42	5.49	15.68

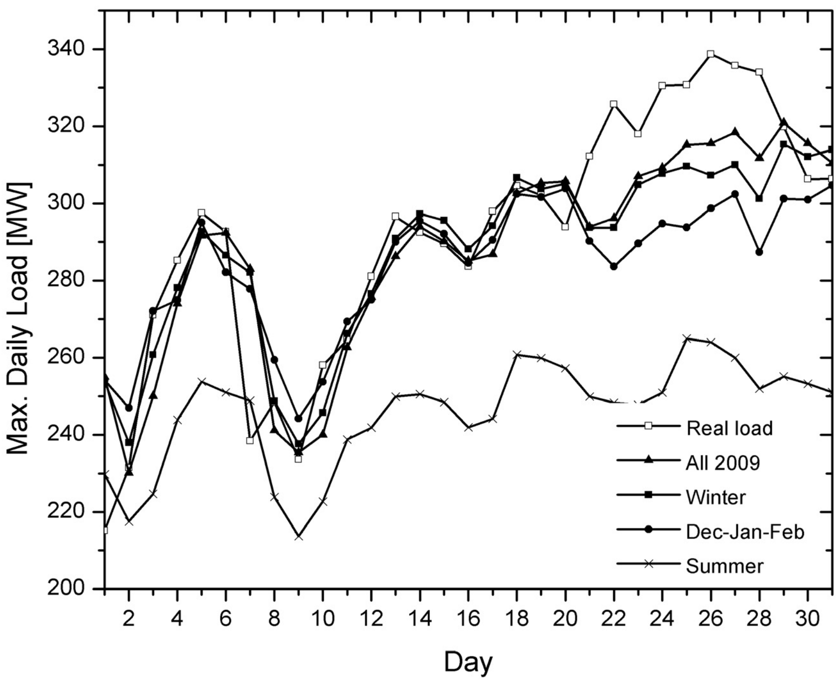

Fig. 5 – *Estimated and real max. daily load demand in Jan.* 2010.

Fig. 5 shows the actual and predicted loads for the month of January 2010 obtained with different training sets. It is obvious that the best prediction is obtained with the training set based on the data of the whole year, but the prediction obtained with data from the winter period is nearly as well.

Table 2
Estimated and real max. load demand per day obtained from training set for 2009.

Day	Real load [MW]	Estimated load [MW]	Relative error [%]
1	215.1	255.0	18.52
2	231.5	230.2	0.56
3	271.0	250.0	7.75
4	285.2	274.0	3.92
5	297.6	291.7	1.98
6	292.7	292.4	0.09
7	238.4	283.0	18.69
8	248.5	241.2	2.93
9	233.6	235.5	0.81
10	258.0	240.1	6.96
11	264.6	262.7	0.73
12	281.1	275.7	1.90
13	296.6	286.3	3.47
14	292.5	294.0	0.51
15	289.6	290.2	0.21
16	283.6	285.1	0.52
17	297.9	286.9	3.72
18	304.6	302.6	0.64
19	302.1	305.3	1.06
20	293.9	305.8	4.03
21	312.2	294.0	5.85
22	325.8	296.2	9.08
23	318.0	307.1	3.43
24	330.5	309.3	6.43
25	330.8	315.2	4.71
26	338.7	315.6	6.83
27	335.8	318.5	5.15
28	334.0	311.8	6.66
29	319.8	321.0	0.36
30	306.3	315.6	3.03
31	306.4	310.5	1.33

Table 1 shows the prediction errors for different training sets. The first row in the table represents the training set, while the second row shows the MAPE. The table reveals the lowest prediction error if the model uses a set of data from the previous year. In the case of training the model with data related to the winter period (from October to March) it is obtained the approximate accuracy as in the previous case. The slightly higher value of error is when the training

set consists of data for the three winter months (December, January and February), while the error was significantly higher when the model was formed on the basis of data from the summer period.

From Fig. 5 and the results given in **Table 2**. it can be observed that the highest prediction error occurs in days 1 and 7, which can be explained that these days are holidays (New Year and Christmas). In the initial stage of feature selection it was planned that the holidays are among them. This was later dropped because the number of holidays in a year is too small to form a suitable model, i.e. model should therefore treat this information as "noise".

5 Conclusion

Proposed SVM model successfully solves the problem of predicting maximum daily electrical load for a period of a month. It is formed a number of models based on different training sets, using data related to different time periods during the year. It is shown that the choice of data to form models significantly affects the quality of predictions. By selection of appropriate training set, good results can be achieved, as demonstrated in the paper. With the aim of achieving greater accuracy in forecasting the plan is to do more detailed analysis of choice of features to form the model.

6 Acknowledgment

The work presented here was supported by the Serbian Ministry of Education and Science (project III44006).

7 References

[1] A. Jain, B. Satish: Clustering based Short Term Load Forecasting using Support Vector Machines, IEEE Power Tech Conference, Bucharest, Romania, June/July 2009, pp. 1 – 8.

[2] D.O. Arroyo, M. Skov, Q. Huynh: Accurate Electricity Load Forecasting with Artificial Neural Networks, International Conference on Computational Intelligence for Modeling, Control and Automation, Vienna, Austria, Nov. 2005, pp. 94 – 99.

[3] V. Vapnik: The Nature of Statistical Learning Theory, Springer, NY, 2000.

[4] S.R. Gunn: Support Vector Machines for Classification and Regression, Technical Report, Image Speech and Intelligent Systems Research Group, University of Southampton, 1998.

[5] V. Cherkassky, Y. Ma: Selection of Meta-parameters for Support Vector Regression, International Conference on Artificial Neural Networks, Madrid, Spain, Aug. 2002, pp. 687 – 693.

[6] D. Basak, S. Pal, D.C. Patranabis: Support Vector Regression, Neural Information Processing – Letters and Reviews, Vol. 11, No. 10, Oct. 2007, pp. 203 – 224.

[7] A.J. Smola, B. Schölkopf: A Tutorial on Support Vector Regression, Statistics and Computing, Vol. 14, No. 3, Aug. 2004, pp. 199 – 222.

[8] N. Amjady: Short-term Hourly Load Forecasting using Time-series Modeling with Peak Load Estimation Capability, IEEE Transactions on Power Systems, Vol. 16, No. 3, Aug. 2001, pp. 498 – 505.

[9] C.W. Hsu, C.C. Chang, C.J. Lin: A Practical Guide to Support Vector Classification, Department of Computer Science, National Taiwan University, Taiwan.

[10] P. Bunnoon, K. Chalermyanont, C. Limsakul: A Computing Model of Artificial Intelligent Approaches to Mid-term Load Forecasting: a State-of-the-art Survey for the Researcher, International Journal of Engineering Technology, Vol. 2, No. 1, Feb. 2010, pp. 94 – 100.