International Journal on Future Revolution in Computer Science & Communication Engineering ISSN: 2454-4248
Volume: 4 Issue: 2 83 – 88

Application of Core Vector Machine for Prediction of Day-Ahead Electricity Prices

R.Gomathi
Department of Electrical Engineering,
Annamalai University,
Chidambaram, India,
e-mail: gomesanthi@gmail.com

S.Ganapathy
Department of Electrical Engineering,
Annamalai University,
Chidambaram, India,
e-mail: ganapathy1967@gmail.com

Abstract— This paper presents Core Vector Machine (CVM) applied for short term electricity price forecasting in an Indian energy market. The accuracy in electricity price forecasting is very crucial for the power producers and consumers. With the accurate price forecasting, power producers can maximize their profit and manage short term operation. Meanwhile, consumers can maximize their utilities efficiently. The objective of this research is to develop models for day-ahead price forecasting using CVM during various seasons. A feature selection technique is used along with the CVM to reduce the variables for accurate price forecasting. Simulation results reveal that the CVM along with feature selection gives better results when compared with other machine learning techniques.

Keywords- Core Vector Machine, Electricity Price Forecasting, Feature Selection, Relief Algorithm, Support Vector Machine

*****

I. INTRODUCTION

Short term electricity price forecasting has gained importance as market participants in the energy industry are exposed to the risk of electricity price variations or try to profit from volatile prices. Short-term forecast cover the period of few days ahead [1]. However, electricity price is a very special commodity as it has a feature that the electricity is not stored unlike other commodities. As a result, electricity price forecasting is considered as an important task under the new environment of deregulated power system [2]. Currently, market participants use different instruments to control and minimize the risks as a result of the volatility introduced by the electricity prices. If the electricity market reference price can be predicted properly, generators and retailers can reduce their risks and maximize their outcomes further [3].

The objective of a price forecasting model is to visualize future electricity market dynamics to assist in present decision making process. Price forecasting accuracy is not as stringent as that of load forecasting. There are various electricity markets around the world using different types of time series models. However, each market uses a suitable forecast model for its own method of functioning. As a result, it is necessary to develop an accurate forecasting model relevant to a particular electricity market [4]. Driven by the importance of the future prices and the complexities involved in determining them, detailed modeling and forecasting of electricity prices has become a major research field in electrical engineering [5]. Forecasting electricity price is a complex task because price series is a non stationary, high frequency and high volatile series with non constant mean, variance and significant outliers as well as calendar effects [6]. Furthermore, electricity prices can rise by tens of or even hundreds of times of its normal value showing one of the greatest volatilities among all the commodities.

Electricity price forecasting has become one of the main endeavors for researchers and practitioners in energy market. Several techniques have been tried for forecasting prices. The majority of them adopt tool from time series models. Two time series models have been proposed and compared with each other in [7]: Dynamic Regression (DR) and Transfer Function (TF) models. In addition, Autoregressive Integrated Moving Average (ARIMA) analysis method has been applied on next-day price forecasting in [8], while a method combining ARIMA and Wavelets has been introduced in [9]. Furthermore, two methods which adopted the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model and Weighted Nearest Neighborhood (WNN) technique respectively have been presented in [10, 11]. Several studies have focused on using Artificial Neural Network (ANN) for price forecasting. ANN was used either as standalone tools [12, 13] or in combination of Fuzzy Logic (FL) [14, 15] and Wavelets [16]. A method which combined Bayesian Expert with Support Vector Machine (SVM) has been introduced in [17]. The performance of each of these existing techniques is highly problem dependent and hence lacks generalization ability. Moreover, because of the nature of the input features used, these methods are found to be incapable of quickly predicting the price. So, in this paper, a new methodology for electricity price forecasting is introduced. The methodology implements Core Vector Machine (CVM) for making day-ahead prediction. CVM is a technique used to overcome the problem of handling large datasets. This method is proven to give better forecast accuracy with lower complexity during optimization, training and forecasting procedures.

II. METHODOLOGY OF CVM

CVM has initially been proposed in [18, 19]. Extensions to the task of support vector regression and classification with class dependent penalties have been proposed in [20, 21]. For solving quadratic programming, SVM learning may become infeasible in large scale settings when datasets comprise several hundred thousand data. The practical algorithms for SVM learning do not solve optimally but impose a tolerance parameter on the Karsh-Kuhn-Tucker conditions, Tsang et al [19] have proposed to reformulate as an equivalent Minimum Enclosing Ball (MEB) problem whose solution can be approximated by means of a fast iterative algorithm using the concept of core sets [18].

Given a set of points i.e., $x_i \varepsilon s$, the MEB is defined as the smallest ball which contains all points. Let r denote the radius and c the center of a ball, the problem of finding an MEB in the feature space can be stated as follows (1):

$$\min_{r,c} r^2$$
$$\|\varphi(x_i)-c\| \leq r^2 \quad (1)$$

The corresponding dual is given as:

$$\max_\alpha \sum_{i,j=1}^{M} \alpha_i \alpha_j k(x_i, x_j) - \sum_{i=1}^{M} \alpha_i k \quad (2)$$

Such that

If, $K(x_i, x_i) = k$, a constant (3)

The second term is discarded in the objective to obtain the final MEB problem (4).

$$\max_\alpha \sum_{i,j=1}^{M} \alpha_i \alpha_j k(x_i, x_j)$$
$$\sum_{i=1}^{M} \alpha_i = 1, \alpha_i \geq 0, \forall_i = 1, \dots M \quad (4)$$

Equation (3) holds good for kernel functions, including the Radial Basis Function kernel. However, a generalization of the CVM [21] does not require this constraint anymore and enables arbitrary kernels. Using a squared error loss function, the dual (5) is obtained with $\delta_{i,j}$ being the Kronecker delta:

$$\max_\alpha = \sum_{i,j=1}^{M} \alpha_i \alpha_j \left(y_i y_j k(x_i, x_j) + y_i y_j + \frac{\delta}{c}\right. \quad (5)$$
$$\sum_{i=1}^{M} \alpha_i = 1, \alpha_i \geq 0, \forall_i = 1, \dots M$$

Now to obtain (4), set:

$$\breve{K}(x_i x_j) = \left(y_i y_j k(x_i, x_j) + y_i y_j + \frac{\delta_{ij}}{c}\right) \quad (6)$$

The computational advantage of solving the MEB problem is with an approximation algorithm from the concept of core sets. Given a set of points $x_i \varepsilon s$, a subset Q c S is a core set of S if an expansion by a factor $(1+\varepsilon)$ of its MEB contains S [18], where ε is a small positive number. Let $B_t(c_t, r_t)$ denote the MEB of the core set Q at iteration t. then the algorithm adds to Q the furthest point outside the ball $B(c_t, (1+\varepsilon)r_t)$. This step is repeated until all points in S are covered by $B(c_t, (1+\varepsilon)r_t)$.
CVM efficiency on large data sets can be attributed to the fact that the size of the final core set depends only on ε but not on M or N [19].
The calculation of class predictions using the CVM differs from the modified kernel Ǩ, which encodes label information y_i, namely:

$$y(x) = \text{sign}\left(\sum_{i \varepsilon Q} \alpha_i \breve{K}(x_i, x) + b\right) \quad (7)$$

The sign function is removed to obtain a continuous prediction which represents the confidence of the classifier [22].

Algorithm [23]

1. Initialize S_0, C_0 and R_0
2. If none of the training point r falls outside the $(1+\varepsilon)$ ball $B(c_t, (1+\varepsilon)r_t)$, the computation terminates. Otherwise, find z furthest away from c_t, and then $S_{t+1} = S_t \cup \{z\}$
3. Find the new MEB (S_{t+1}) and use to calculate new parameters C_{t+1} and R_{t+1}, $t = t+1$ and go back to step 2.

1. From the above steps, all the points added to the core set will be called core vectors. Despite its simplicity, CVM greatly reduces the computation complexity compared with the SVM algorithm.

A. Importance of feature selection

Feature selection is one of the important and frequently used techniques in data reduction or preprocessing for data mining. There are a number of advantages of feature selection: it reduces the number of features, removes irrelevant, redundant or noisy data, reduces the computational cost, speeds up a data mining algorithm and improves the classification accuracy [24]. Feature selection is an important preprocessing step to machine learning. It is a process that selects a subset of original features. The optimality of a feature subset is measured by an evaluation criterion [25]. In this work, a filter approach i.e., Relief Algorithm [26] has been used.

B. Relief Algorithm

Relief is considered as an efficient and easy to use feature selection method among the other techniques. Relief Algorithm assigns weight to candidate input features, which is inspired by instance based learning algorithms. It first randomly chooses a number of training samples from training set [27]. For each selected samples, it finds the nearest hit and nearest miss instances based on Euclidean distance measure:

$$d_{rs} = \sqrt{x_r^2 + x_s^2 - 2x_r x_s} \quad (8)$$

ISSN: 2454-4248
83 – 88

where x_r and x_s are two vectors representing samples r and s respectively and d_{rs} is Euclidean distance between them. Nearest hit and nearest miss are two samples having minimum Euclidean distance of the same and opposite classes respectively. The relief updates the weights of the features that are initialized to zero in the beginning, based on an intuitive idea that a feature is more relevant if it distinguishes between a sample and its nearest miss are less relevant if it distinguishes between a sample and its nearest hit. In the relief algorithm, the weight of i^{th} feature w_i is updated according to the following equation:

$$w_i = w_i + |x^{(i)} - NM^{(i)}(x)| - |x^{(i)} - NH^{(i)}(x)| \qquad (9)$$

where $x^{(i)}$ indicates i^{th} feature of the selected samples x; $NM^{(i)}$ (x) and $NH^{(1)}$ (x) represent i^{th} feature of the nearest miss and nearest hit of the sample x, respective; i is number of candidate input features. In this way, weights of all candidate inputs are updated based on the selected samples x. This cycle is repeated for all randomly selected samples and then the candidate inputs are ranked according to the finally obtained weight values.

After exhausting all instances in the chosen sample, variables with high weighting values are added in the feature vector. In this work, the variables have been chosen around 20% having highest weights as decision making features.

The weight of i^{th} feature, based on the k nearest neighborhoods, can be evaluated as follows:

$$W_i = \frac{\sum_{t=1}^{k} WNM_{ti}}{\sum_{t=1}^{k} WNH_{ti}} \qquad (10)$$

To calculated the weight value w_i, equation (10) is replaced by the ratio between two summations of all distances to the nearest misses and nearest hits respectively.

$$\left.\begin{aligned} WNH_{1i} &= \frac{1}{N}\sum_{l=1}^{N} WS_l\,|DNH_{1(i,l)}| \\ WNH_{ki} &= \frac{1}{N}\sum_{l=1}^{N} WS_l\,|DNH_{k(i,l)}| \end{aligned}\right\} \qquad (11)$$

$$\left.\begin{aligned} WNM_{1i} &= \frac{1}{N}\sum_{l=1}^{N} WS_l\,|DNM_{1(i,l)}| \\ WNM_{ki} &= \frac{1}{N}\sum_{l=1}^{N} WS_l\,|DNH_{k(i,l)}| \end{aligned}\right\} \qquad (12)$$

Where the subscript (i,l) indicates the element located in the row i and column l of the respective matrix . WS_l represent weight of sample l, since the training patterns may have different information values. In (11) and (12), WNH_{ki} and WNM_{ki} measure the effect of i^{th} feature in distinguishing the samples from their nearest hits and misses of order k respectively

The weight values obtained from w_i are normalized with respect to their maximum such that all the weights be in the range of [0,1]. Then, the candidate inputs are ranked based on their normalized weight values. The candidate inputs with the normalized weight larger than a threshold are selected by a relief algorithm and the other candidates are filtered out. The threshold is determined by the cross validation technique.

Finally, the parameter of k i.e. the number of neighborhoods should be determined. The large values of k increase computation burden at the same time further neighborhoods are included which may be ineffective and never misleading. The k parameter can be obtained from the following relation:

$$k = \mathrm{Round}(\log_2 N) \qquad (13)$$

N is the number of samples. Round is a function that rounds the real number to the closest integer value. From k parameter, the value obtained from (13) leads to the best results for feature selection, especially for electricity price forecasting [28].

III. Implementation

After ranking the input features by the Relief Algorithm, the training instances can be constructed by the historical data. The day-ahead price forecast is reached via recursion. In this case, when price of an hour is forecasted, it is used as P_i for the price prediction of the next hour and this cycle is repeated until the price of the next 168 hours are predicted. In the day-ahead price forecast, the price data is updated once every day.

In this paper, the prices of the previous 42 days are used in the training period, which result in 1008 training instances. The data for the electricity prices have been taken from the daily trading reports of an Indian market that are presented on an hourly basis, customized on yearly reports. The test period of the seasons is Winter: October 26- November 1, Spring: February 5-11, Summer: June 22-28, Rainy: September 21-27. As the test period is of one week and for a day it is 24 hours, the total for a week is 168 hours. The total 168 variables are reduced by using the Relief Algorithm. The proposed work is examined on an Indian market of the year 2016 and the results show the effectiveness of the proposed CVM approach.

A. *Performance evaluation*

The error criterion, namely, Mean Absolute Percentage Error (MAPE) is used to evaluate the accuracy of forecasting models [4].

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N} \frac{P_i - P_i'}{P_i} \times 100 \qquad (14)$$

Where N is the time periods, P_i and P_i' are the actual and forecasted electricity prices respectively.

International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 4 Issue: 2 ISSN: 2454-4248
83 – 88

B. Results and discussion

Table 1 represents the input features, which are selected using Relief Algorithm. Table 2 presents the values to evaluate the accuracy of the proposed CVM and CVM-Relief Algorithm models in forecasting the electricity prices on an Indian Energy Exchange market. The Indian Energy Exchange electricity market has an average value of 7.66% and 6.47% obtained using CVM and CVM-Relief models. It is clear from Table 2 that the accuracy increases with decrease in MAPE when feature selection is incorporated in the CVM model.

TABLE 1: FEATURE SELECTION PROCESS

Case Study	Indian Market
No. of pattern variables	168
No. of features selected	77
Dimensionality reduction	46%
Selected features	P_{56}, P_{71}, P_{129}, P_{127}, P_{16}, P_{14}, P_{163}, P_{15}, P_{45}, P_{94}, P_{86}, P_5, P_{128}, P_{35}, P_{124}, P_{123}, P_{92}, P_{30}, P_{120}, P_{57}, P_{46}, P_6, P_{38}, P_{58}, P_{122}, P_{126}, P_{44}, P_{95}, P_{54}, P_{39}, P_{25}, P_{12}, P_{29}, P_{31}, P_{59}, P_{34}, P_{20}, P_{32}, P_{121}, P_{78}, P_{42}, P_{36}, P_7, P_{13}, P_1, P_{50}, P_{156}, P_{18}, P_2, P_9, P_{83}, P_{22}, P_{37}, P_{11}, P_{23}, P_{60}, P_{125}, P_{40}, P_{91}, P_{10}, P_{144}, P_{43}, P_{47}, P_{143}, P_{73}, P_{41}, P_{112}, P_{82}, P_{76}, P_{21}, P_{52}, P_{69}, P_8, P_{66}, P_{62}, P_4, P_{155}

TABLE 2 STATISTICAL ANALYSIS FOR THE FOUR WEEKS OF INDIAN ENERGY EXCHANGE MARKET USING CVM

Market	Forecast Week	MAPE% (CVM)	MAPE% (CVM-Relief)
	WINTER		
	SPRING	9.49	7.87
	SUMMER	7.04	6.34
Indian	RAINY	7.09	5.78
	AVERAGE	7.02	5.92
		7.66	6.47

Table 3 shows the comparison between machine learning techniques of CVM and SVM with feature selection. CVM gives lower MAPE compared to SVM.

TABLE 3 COMPARATIVE MAPE RESULTS

Indian Market		
SEASONS	**SVM**	**CVM**
WINTER	8.08	7.87
SPRING	7.54	6.34
SUMMER	7.04	5.78
RAINY	7.45	5.92

C. Forecasted MCP versus Actual MCP for the Indian Market

Here, Figures 1-4 show the four seasons of the actual market clearing price and forecasted market clearing price using CVM-Relief. The X-axis represents the hours for the week and Y-axis represents the prices of Indian market. It is observed from the figures 1-4 that the price prediction indicates the superior capability of the proposed method used for converting ill behaved price series into a set of better behavior series. The other reason for superiority of this method may be attributed to the suitable selection of input variables for each sub series using Relief algorithm. Moreover, this proposed method has its ability for capturing the features of linearity and non linearity associated with electricity prices. Therefore, the proposed method gives better prediction [4].

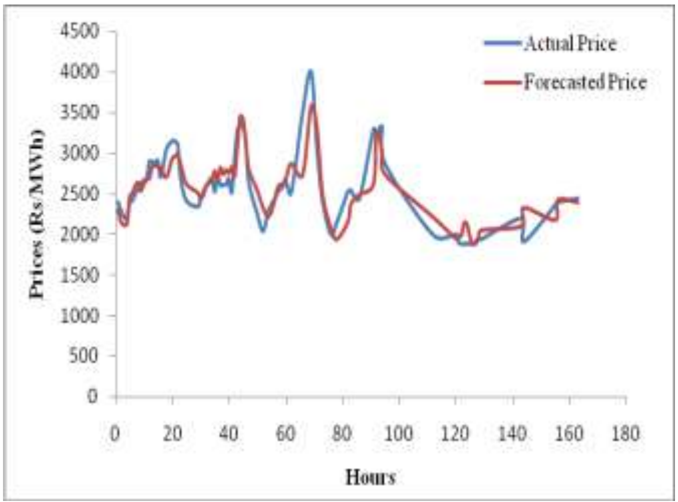

Figure 1. Forecasted Market Clearing Price for a Winter Week (October 26 to November1)

ISSN: 2454-4248
83 – 88

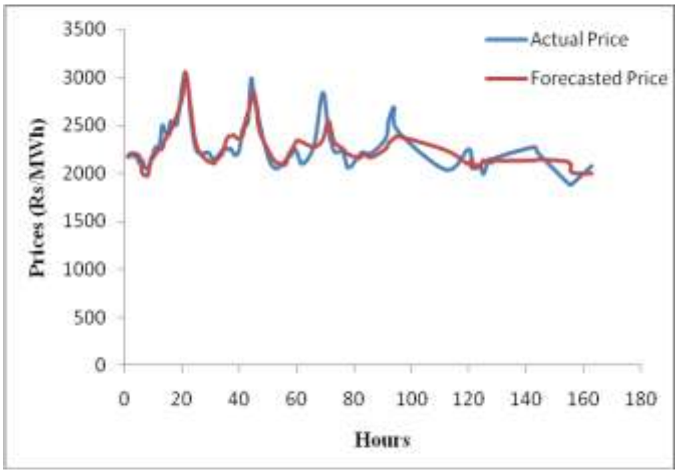

Figure 2. Forecasted Market Clearing Price for a Spring Week (February 5 to February 11)

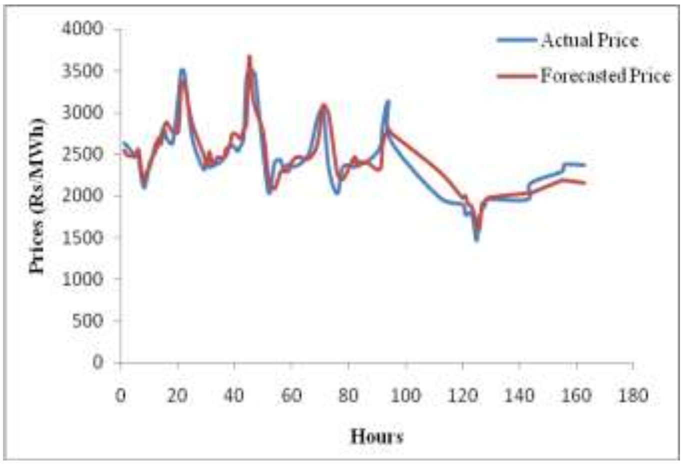

Figure 3. Forecasted Market Clearing Price for a Rainy Week (September 21 to September 27)

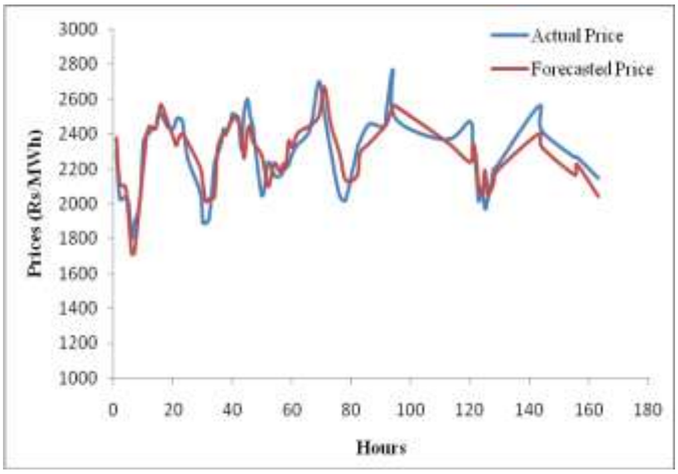

Figure 4. Forecasted Market Clearing Price for a Summer Week (June 22 to June 28)

IV. Conclusion

Accurate price forecasting is very important for electric utilities in a competitive environment created by the electric industry deregulation. Historical information of electricity price in forecasting the day-ahead price is presented. A new price forecasting method based on CVM and CVM combined with Relief Algorithm models is proposed in this paper. The results from the comparisons clearly show that the CVM-Relief Algorithm is far more accurate than the other forecast methods. The proposed method has been examined on an Indian market in the year 2016. The experimental results indicate that the proposed method can provide an accurate and effective forecasting.

References

[1] G. Aarti, C. Pankaj, C. Sparsh, “Short Term Electricity Price Forecasting using ANN and Fuzzy Logic under Deregulated Environment”, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 2013; 2: 3852-3858.

[2] H. Mori, A.Akira, “Data Mining of Electricity Price Forecasting with Regression Tree and Normalized Radial Basis Function Network”, in the Proceedings of IEEE International Conference on Systems, Man and Cybernetics, Montreal, Canada, 2007; 3743-3748.

[3] “Self-Adaptive Radial Basis Function Neural Network for Short Term Electricity Price Forecasting ”, IET Generation, Transmission and Distribution, 2009; 3: 325-335.

[4] S.E. Peter, I. Jacob Raglend, “Sequential Wavelet-ANN with Embedded ANN-PSO Hybrid Electricity Price Forecasting Model for Indian Energy Exchange”, Neural Computing and Applications, 2017; 28: 2277-2292.

[5] A.S. Nitin, P.K. Bijaya, “A Hybrid Wavelet-ELM based Short Term Price Forecasting for Electricity Market”, International Journal of Electrical Power and Energy Systems, 2014; 55: 41-50.

[6] W.A.R. Intan Azmira, T.K.A. Rahman, A. Arfah, “Short Term Electricity Price Forecasting using Neural Network”, in the Proceedings of Fourth International Conference on Computing and Information, ICOCI 2013, Sarawak, Malaysia, 2013; 28-30.

[7] F.J. Nogales, J. Contreras, A.J. Conejo, R. Espinola, “Forecasting Next-Day Electricity Prices by Time Series Models”, IEEE Transactions on Power Systems, 2002; 17: 342-348.

[8] F.J. Nogales, J. Contreras, A.J. Conejo, R. Espinola, “ARIMA Models to Predict Next-Day Electricity Prices” , IEEE Transactions on Power Systems, 2003; 18: 1014-1020.

[9] A.J.Conejo, “Day-Ahead Electricity Price Forecasting using the Wavelet Transform and ARIMA Models”, IEEE Transactions on Power Systems, 2005; 20: 1035-1042.

[10] R.C. Garcia, J. Contreras, A. Marco Van, J.B. C. Garcia, “A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices”, IEEE Transactions on Power Systems, 2005; 20: 867-874.

[11] A.T. Lora, J.M. Riquelme Santos, G. Antonio, J.L.M. Ramos, J.C. Riquelme Santos, "Electricity Market Price Forecasting based on Weighted Nearest Neighbors Techniques", IEEE Transactions on Power Systems, 2007; 22: 1294-1301.

[12] B. R. Szkuta, L.A. Sanabria, T.S. Dillion, "Electricity Price Short-Term Forecasting Using Artificial Neural Network", IEEE Transactions on Power Systems, 1999; 14: 851-857.

[13] H.Y. Yamin, S.M. Shahidehpour, Z. Li, "Adaptive Short-Term Forecasting Using Artificial Neural Network in the Restructured Power Markets", International Journal of Electrical Power and Energy Systems, 2004; 26: 571-581.

[14] N. Amjady, "Day-Ahead Price Forecasting of Electricity Markets by a New Fuzzy Neural Network", IEEE Transactions on Power Systems, 2006; 21: 887-896, 2006.

[15] Y.Y.Hong, C.F. Lee, "A Neuro-Fuzzy Price Forecasting Approach in Deregulated Electricity Markets", Electrical Power System Research, 2005; 73: 151-157.

[16] C.A.C. Coello, G.T. Pulido, "A Micro Genetic Algorithm for Multiobjective Optimization", in the Proceedings of the Genetic and Evolutionary Computation Conference, GECCO, Switzerland, 2001; 126-140.

[17] W.Wu, J.Zhou, L. Mo, C. Zhu, "Forecasting Electricity Market Price Spikes based on Bayesian Expert with Support Vector Machine", ADMA, Lecture Notes in Artificial Intelligence LNAI 4093, Berlin: Springer-Verlag, 2006; 205-212.

[18] I.W.Tsang, J.T. Kwok,, and P.M. Cheung, "Core Vector Machine: Fast SVM Training on Very Large Datasets", Journal of Machine Learning Research,2005; 6: 363-392.

[19] I.W.Tsang, J.T. Kwok,, and P.M. Cheung, "Very Large SVM Training using Core Vector Machine", Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics (AISTATS), Barbados, 2005; 349-356.

[20] I.W.Tsang, J.T. Kwok,, and K.T. Lai, "Core Vector Regression for Very Large Regression Problems", in the Proceedings of the Twenty Second International Conference on Machine Learning Bonn, Germany, 2005; 912-919.

[21] I.W.H. Tsang, J.T.Y. Kwok, J.M. Zurada, "Generalized Core Vector Machine", IEEE Transactions on Neural Network, 2006; 17: 1126-1140.

[22] V.N. Vapnik, "The Nature of Statistial Learning Theory" IEEE Transactions on Neural Networks, 1997; 8: 1-15.

[23] X. Li, J. Yu, Z. Jia, H. Wang, "Application of Improved CVM in the Prediction of Algal Blooms in Tolo Harbour", in the Proceedings of Sixteenth IEEE International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, Takamatsu, Japan, 2015; 1-6.

[24] R. Porkodi, "Comparison of Filter based Feature Selection Algorithms: An Overview", International Journal of Innovative Research in Technology and Science, 2014; 2: 108-113.

[25] V.B. Vaghela, K.H. Vandra, N.K. Modi, "Entropy based Feature Selection for Multi-Relational Naïve Bayesian Classifier", Journal of International Technology and Information Management, 2014; 23: 13-26.

[26] S. Francisca Rosario, K. Thangadurai, "Relief: Feature Selection Approach", International Journal of Innovative Research and Development, 2015; 4: 218-224.

[27] K. Igor, "Estimating Attributes: Analysis and Extensions of Relief", in the Proceedings of European Conference on Machine Learning, ECML, Catania, Italy, 1994; 171-182.

[28]N. Amjady, D. Ali, "Day-Ahead Electricity Price Forecasting using the Relief Algorithm and Neural Network", in the Proceedings of IEEE Fifth International Conference on European Electricity Market, Lisboa, Portugal, 2008; 1-7.