Proceedings of the International Conference on Big Data Cloud and Applications
Tetuan, Morocco, May 25 - 26, 2015

Electric load forecasting using hybrid machine learning approach incorporating feature selection

Malek Sarhani
ENSIAS
Mohammed-V University
Rabat, Morocco
malek.sarhani@um5s.net.ma

Abdellatif El Afia
ENSIAS
Mohammed-V University
Rabat, Morocco
abdellatif.elafia@gmail.com

ABSTRACT

Forecasting of future electricity demand is very important for the electric power industry. As influenced by various factors, it has been shown in several publications that machine learning methods are useful for electric load forecasting (ELF). On the one hand, we introduce in this paper the approach of support vector regression (SVR) for ELF. In particular, we use particle swarm optimization (PSO) algorithm to optimize SVR parameters. On the other hand, it is important to determine the irrelevant factors as a preprocessing step for ELF. Our contribution consists of investigating the importance of applying the feature selection approach for removing the irrelevant factors of electric load. The experimental results elucidate the feasibility of applying feature selection without decreasing the performance of the SVR-PSO model for ELF.

I. INTRODUCTION

For developing countries, accurate electric load forecasting (ELF) is an important guide for effective actions of energy policies. Furthermore, accurate models for electric power load forecasting are essential to the operation and planning for several companies. It may have an impact on energy purchasing and generation, load switching, contract evaluation, and infrastructure development. The cost of error is so high that research in forecasting techniques which could help to reduce it in a few percent points would be amply justified.

Load forecasts can be divided into three categories: short-term forecasts which are usually from one hour to one week, medium forecasts which are usually from a week to a year, and long-term forecasts which are longer than a year. Short term forecasting are essential for the control and scheduling of power systems [35]. However, daily load forecasting is a hard task, because it depends not only on the load of the previous days, but also on other facts such as temperature, calendar effect [42].

Nowadays, there are different techniques for calculating forecasts, In one hand, classical statistical foecasting methods such as exponential smoothing (Winter 1960 [43]) or ARIMA models defined by Box and Jenkins (1994) [4] can be used for this purpose. But, with these traditional methods, The construction of ELF model may be difficult due to its uncertain, non-linear, dynamic and complicated characteristics: electric load data present nonlinear data patterns caused by influencing factors such as climate factors, seasonal factors, and so on (Amjady and Keynia 2009) [1]. Thus, methods based on artificial intelligence techniques like artificial neural network (Minsky and Papert in 1969 [25]), genetic algorithms (Goldbergn in 1989 [14]), fuzzy logic (Cox and Earl in 1994 [8]) and support vector machine (SVM) (Vapnik et al. 1997 [38]) can generate better results (Ul Islam 2011 [19]).

In the past few years, various efforts in improving the forecasting accuracy were proposed. Lots of these researchers have tried to apply artificial intelligence techniques to improve forecasting accuracy. The most used method is artificial neural network (ANN). Or, Hu and Zhang (2008) [11] showed that ANN has inherent drawbacks, such as local optimization solution, lack generalization, and uncontrolled convergence. Therefore, support vector machine (SVM), which overcomes some drawbacks of neural networks, was introduced to provide a model with better predictive power to elaborate a more accurate forecast.

Of the influencing factors on ELF which are presented in real dataset, some of them could be redundant or irrelevant. Thus, feature selection (FS) is justified as a first step for ELF. Our contribution in this paper is to investigate the importance of using FS in ELF. The rest of the paper is organized as follows: In the next section, we introduce the concepts related to our forecasting techniques. In the section 3, we outline the related works. Section 4 describes the algorithm and the tools which are used for its implementation and presents the case study used for the evaluation. Section 5 presents the parameters setting. The results are presented in section 6. Finally, we conclude and present perspectives to our work.

II. THE HYBRID MACHINE LEARNING TECHNIQUE

A. Support Vector Machine

The support vector machine (SVM) is a recent tool from the artificial intelligence field which use statistical learning theory. It has been successfully applied to many fields and it recently of increasing interests of researchers: It has been first introduced by Vapnik et al.(1992) [3] and was applied firstly to pattern recognition (classification) problems, recent research has yielded extensions to regression problems, including time series forecasting.

SVM belongs to Kernel methods, which represent a new generation of learning algorithms and utilize techniques from optimization, statistics, and functional analysis in pursuit of maximal generality, flexibility, and performance. SVM applies the structural risk minimization (SRM) principle to minimize an upper bound on the generalization error. SVM could theoretically guarantee to achieve the global optimum.

The main use of SVM is in classification. However, a version of an SVM for regression has been proposed by Vapnik et al. in 1997 [38].

B. Support Vector Regression

This subsection introduces briefly the idea of SVM for the case of regression (SVR). SVR have been successfully employed to solve forecasting problems in many fields, such as financial time series forecasting [20], engineering and software field forecasting [31], and so on.

The basic concept of the SVR model is to nonlinearly map (with function $\varphi(.) : \mathcal{R}^n \rightarrow \mathcal{R}^{n_h}$) the input data (training data set $(x_i, y_i)_{i=1}^N$) into a higher dimensional feature space (which may have infinite dimensions R^{n_h}). Then, the SVR function is shown as follows:

$$f(x) = \omega\varphi(x) + b \tag{1}$$

where $f(x)$ denotes the forecasting values. the coefficients ω and b are estimated by solving the following formulation which aims to minimize the regularized risk function:

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

$$s.t \begin{cases} y_i - (< w, \phi(x_i) > +b) \leq \epsilon + \xi_i \\ (< w, \phi(x_i) > +b) - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \tag{3}$$

The constant C determines the trade off between the flatness of f and the amount up to which deviations larger than ϵ are tolerated. ξ_i denotes the training error above ϵ, whereas ξ_i^* denotes the training error below $-\epsilon$, and n represents the number of samples. SVR avoids underfitting and overfitting of the training data by minimizing the regularization term $\frac{1}{2}||w||^2$ as well as the training error $C\sum_{i=1}^{N}(\xi_i + \xi_i^*)$.

This constrained optimization problem can be solved by the pri- mal Lagrangian form and the KarushKuhnTucker conditions, the dual can be obtained as: Maximize

$$-\frac{1}{2}\sum_{i,j=1}\alpha_i\alpha_j < x_i, x_j > -\epsilon\sum_{i=1}^{l}\alpha_i + \sum_{i=1}^{l} y_i\alpha_i \tag{4}$$

where$\alpha_i = \beta_i - \beta_i*$ and β_i*, β_i are obtained by solving the quadratic program and are the Lagrangian multipliers. After this optimization problem is solved, the parameter vector w in Equation (2) is obtained by:

$$w = \sum_{i=1}^{N}(\beta_i^* - \beta_i)\varphi(X_i) \tag{5}$$

. Finally, the SVR regression function is obtained as the following equation in the dual space

$$f(x) = \sum_{i=1}^{N}(\beta_i^* - \beta_i)K(x_i, x) + b \tag{6}$$

where $K(x_i, x_j)$ is called the kernel function: The value of the kernel equals the inner product of two vectors x_i and x_j in the feature space $\varphi(x_i)$ and $\varphi(x_j)$. The most commonly used kernel functions are the Gaussian radial basis functions (RBF) kernel function, namely $K(x_i, x_j) = exp(-||x_i - x_j||^2/2\sigma^2)$ which is also employed in this study.

C. Particle Swarm Optimization

The parameters that should be optimized include the penalty parameter C, ϵ and ω defined in equation (2). Thus, the choice of the parameters has a heavy impact on the forecasting accuracy. The PSO algorithm is used to seek a better combination of the three parameters in the SVR.

Particle swarm optimization (PSO) was originally designed by Kennedy and Eberhart in 1995 [22]. The technique simulates the moving of social behaviour among individuals (particles) through a multi-dimensional search space, each particle represents a single intersection of all search dimensions.

In PSO, Each particle i has two vectors: the velocity vector and the position vector: The particles are updated according to itself previous best position and the entire swarm previous best position. That is, particle i adjusts its velocity ν_i and position x_i in each generation according to the following formula:

$$\begin{cases} \nu^{n+1} = \omega\nu^n + c_1r_1(p^n - x^n) + c_2.r_2.(p_g^n - x^n) \\ x^{n+1} = x^n + \beta.\nu^n \end{cases} \tag{7}$$

where ν^n and x^n are the current velocity and position of the particle. p^n represents the best previous position of particle i. p_g^n represents the best position among all particles in the population. r_1 and r_2 are two independently uniformly distributed random variables with a range.

Nowadays, PSO has gained much attention and wide applications in solving continuous non linear optimization problems due to the simple concept, easy implementation and quick convergence. [5].

D. Feature selection

Feature selection (FS), also known as variable selection or attribute selection, aims at identifying the most relevant input variables within a dataset. It may improve the performance of the predictors by eliminating irrelevant inputs, achieves data reduction for accelerated training and increases computational efficiency [34]. It is usually utilized to identify a subset where the meanings of variables are important.

Most feature selection algorithms perform a search through the space of feature subsets. There are some characteristics which affect the nature of the search: The most important are the search organization (heuristic strategies are generally more feasible and adaptable for this problem), and the evaluator (we can distinguish two major families of methods: Filter and wrapper).

Moreover, selecting the key variables is crucial in constructing the energy load forecasting model. Furthermore, according to Lu (2014) [23], the major disadvantage of SVR is that it cannot select important variables from many predictor variables.

III. Literature review

Related work for this research includes the use SVM and SVR for Electric load forecasting (ELF) in general. Particularly, we focus on works which used the hybrid model SVR-PSO for load forecasting and others which interest in the selection of relevant attributes.

That is, SVM and SVR was being applied to ELF. For instance, Mohandes (2002) [26] applied the method of SVM for short-term ELF. The obtained results for this paper indicate that SVM outperforms the autoregressive method. Also, Wang et al. (2009) [41] presented a ϵ -SVR model considering seasonalproportions based on development tendencies from history data. Since electric load data are non-linear in relation and complex, many studies tend to hybridize SVR with other methods, Elattar et al. (2010) [12] proposed an approach for solving the load forecasting problem which combines the and locally weighted regression. Then, he employed the weighted distance algorithm that uses the Mahalanobis distance to optimize the weighting function's bandwidth. In the study of Ogcu et al. (2012) [30], SVR and ANN models were employed to develop the best model for predicting electricity output. Che (2012) et al.[6] presented an adaptive fuzzy combination model based on the self-organizing map (SOM), the SVR and the fuzzy inference method. Furthermore, several algorithms have been proposed to optimize SVR parameters. Hong et al. (2011) [18] introduced the application of Chaotic Immune Algorithm for optimizing SVR parameters and investigate its feasibility for ELF. Zhang et al. (2012) [45] investigated the potentiality of a hybrid algorithm which combine chaotic genetic algorithm and simulated annealing algorithm for optimizing SVR model and improving load forecasting accurate performance. Another hybrid forecasting model using differential evolution algorithm to determine the parameters in SVR model was proposed by Wang et al. (2012) [40] for forecasting the annual electric load. Aung et al. (2012) [2] adopted the least-squares support vector regression technique incorporated with online learning to forecast the peak load of a particular consumer entity in the smart grid for a future time unit.

Furthermore, the SVR-PSO applied for ELF: Hong (2009) [17] elucidated the feasibility of applying chaotic particle swarm optimization (CPSO) algorithm to choose the suitable parameter combination for a SVR model in forecasting of electric load. Duan et al. (2011) [10] proposed a combined method for the short-term load forecasting of electric power systems based on the Fuzzy-c-means (FCM) clustering, PSO and SVR techniques. The SVR-PSO has been used for forecasting in other fields. For instance, Anandhi et al. (2013) [37] presented an SVR based prediction model appropriately tuned can outperform other more complex models. Specially, evaluated results show that proposed SVM regression with PSO approach gave improved accuracy. This approach which combines SVR and PSO was presented also to traffic safety forecasting in the paper of Gang and Zhuping (2011) [13].

Tu et al. (2007) [36] performed feature selection with PSO and used SVM to evaluate the fitness value. He et al. (2008) [16] used Genetic algorithm for feature selection which lead to reduce input space. Nguyen and Torre (2010) [28] have presented a method for jointly learn weights and parameters of the SVM model. Crone and Kourentzes (2010) [9] proposed an iterative neural filter is proposed for feature evaluation to automatically identify the frequency of the time series.

In their paper, Vieira et al. (2013) [39] proposed a binary PSO algorithm for feature selection in parallel with optimizing the SVM parameters. Lu (2014) [23] used Multivariable Adaptive Regression Splines (MARS) for selecting input variables and then construct a sales forecasting model with SVR. Shahrabi et al. (2013) [33] presented an approach which used k-means clustering for reducing the dimension of the data space, and then used genetic expert system for forecasting tourism demand.

Niu et Guo (2009) [29] proposed a method which uses simulated annealing to improve the global searching capacity of the PSO for the purpose of optimizing SVR parameters and selecting its input features and then applied it to short term load forecasting. Karimi (2012) [21] proposed a feature selection technique composed of Modified Relief and Mutual Information and then forecast electric load with a training mechanism. Yadav et al. (2014) [44] used Weka software in order to select the most relevant input parameters for solar radiation prediction models.

IV. The proposed approach for ELF

A. The SVR-PSO model

Resolving the SVR dual problem is often troublesome. Despite an exhaustive search method could be used to tune this, it suffers from the main drawbacks of being very time-consuming and lacking of a guarantee of convergence to the globally optimal solution. Compared to genetic algorithms (GA), the PSO method can efficiently find optimal or near-optimal solutions in large search spaces. Furthermore, Lu and Geng (2011) [24] showed that the PSO-SVR model is superior to GA-SVR model in the running efficiency and predictive accuracy. Thus, we adopt PSO for optimal parameter selection of SVR in order to improve the accuracy and runtime efficiency of learning procedure of SVR-PSO. Our SVR-PSO algorithm for ELF can be defined as following:

Initialize $Pop(\alpha_i)$, Initialize σ, ϵ, C
while $t \leq t_{max}$ **do**
  for i = 0 to n **do**
    Compute $f_{\alpha\, i}$ according to Eq. (4)
    Update $V_{\alpha\, i}$ and $X_{\alpha\, i}$ according to Eq. (7)
    if $f_{\alpha\, i} \leq f_{best\ \alpha\, i}$ **then**
      $f_{best\ \alpha i} -> f_{\alpha\, i}$
      $X_{best\ \alpha i} -> X_{\alpha\, i}$
    else $\{N$ is odd$\}$
      $N \leftarrow N - 1$
    end if
  end for

end while

B. SVR-PSO with feature selection

As mentioned in the previous section, the SVR-PSO model is useful for electric load forecasting (ELF). Therefore, we apply it to the electric load forecasting. Moreover, we use feature selection to remove irrelevant attributes as illustrated below, The procedure used in this paper is summarized in Figure 2.

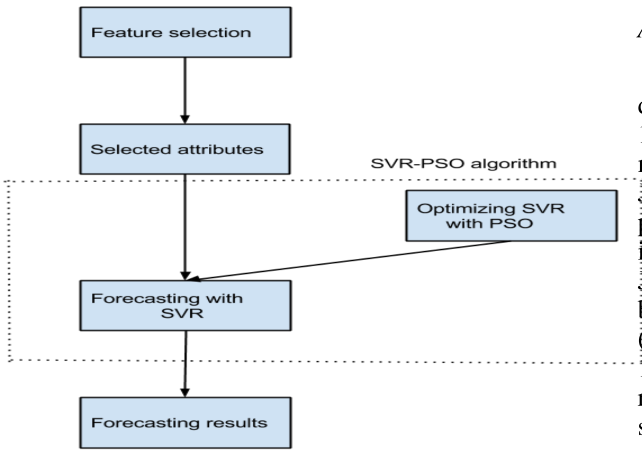

Fig. 2. SVR-PSO with feature selection

At the best of our knowledge, this hybrid SVR-PSO model combined with feature selection hasn't been yet applied for ELF. This idea has been investigated in the paper of [32], but the results of the forecasting model aren't enough to validate this approach.

V. EXPERIMENTS SETUP

A. Tools

To perform feature selection, Weka software was used. Weka is a collection of machine learning algorithms for data mining tasks.

As mentioned in section II.D, feature selection has two principal characteristics. In this study, we used Correlation-based Feature Subset Selection (CFS) as an attribute evaluator (see Hall 1998 [15]) and Particle Swarm Optimization for search method (see Moraglio et al. 2009 [27]).

Forecasting with SVR involves normalization of data within the range of [0,1]. Also, the SVR-PSO method with and without FS have been executed on the same platform: Intel Core i5 PC, 1.8 GHz with 4 GB of RAM under Ubuntu 14.04 operating system.

B. Comparison measurement

The experimental data should be divided into two subsets: the training set and the testing set. The forecasting accuracy is measured in the testing set by two criteria which are Mean Absolute Percentage Error (MAPE) and Mean Square Error (MSE). The MAPE and MSE are given by the following equations:

$$MAPE = 100.\frac{\sum |(prediction - real)/real|}{n} \tag{8}$$

$$MSE = \frac{(prediction - real)^2}{n} \tag{9}$$

Where n is the number of instance of the testing set.

VI. EXPERIMENTAL RESULTS

A. Experiment 1:

First of all, the paper takes The historical electricity load dataset used in the EUNITE competition [7] from January 1, 1997 to December 31, 1998 as shown in Table I. The maximum daily values of the electrical load for the 31 days of January 1999 are to be forecasted using the given data for the preceding two years. Given load and some other information in 1997-1998, the task is to predict daily maximum load in January 1999. This dataset contains 16 features. As described by Chen et al. [7], these features belong to three categories (calendar attributes, temperature, past load demand). Features 1-7 correspond to the seven days of the week. Feature 8 is related to temperature. Features 10-16 are loads of the previous seven days.

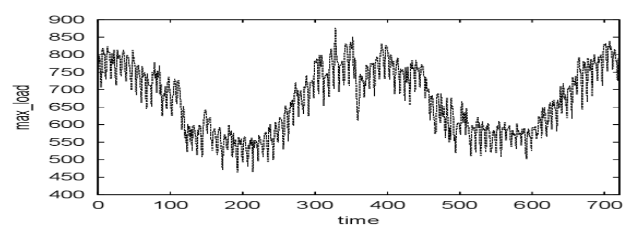

Fig. 3. Electric load history for experiment 1

, we apply feature selection for the EUNITE competition dataset:

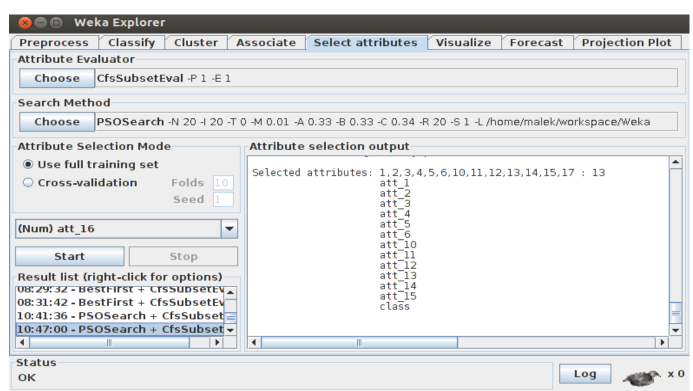

Fig. 4. The selected attributes

Figure 4. shows the feature selection result using Weka. It can be found that thirteen attributes are selected as important features (The attributes 7,8 and 9 are eliminated). Thus, these

features were chosen for being applied for building the SVR-PSO model of ELF.

Below, we will apply the SVR-PSO model the EUNITE case study and shows the predicted values against the realist values. The figure 5. presents it for the case of the model without feature selection, while the figure 6. shows it for the case with feature selection. The following table shows different performance measurement obtained:

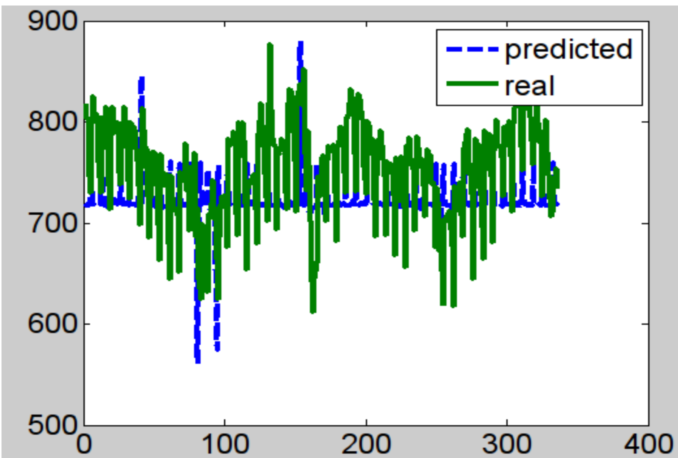

Fig. 5. SVR-PSO without feature selection

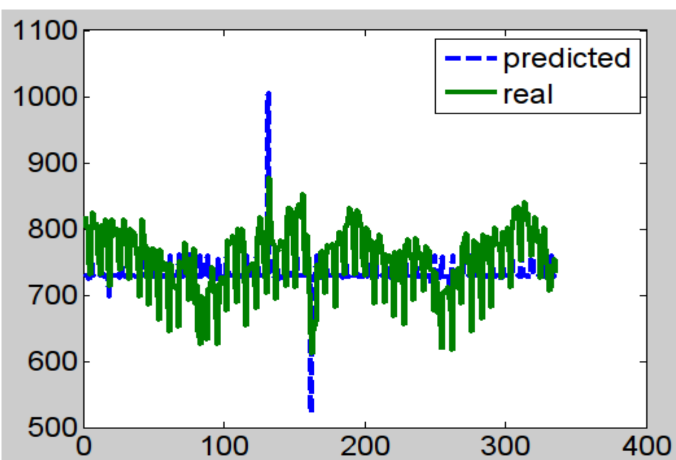

Fig. 6. SVR-PSO with feature selection

	SVR-PSO without FS	SVR-PSO with FS
MAPE	0.0613	**0.0555**
MSE	$3.0275.10^3$	$2.5533.10^3$

Table 1. Comparison of results for experiment 1.

The first column presents the performance of SVR-PSO without feature selection and the second column presents it for SVR-PSO with feature selection.

On one hand, the eliminated attributes in the section VI.A are 7,8 and 9. The attribute 9 wasn't used in the competition as mentioned in V-B. The attribute 7 is Sunday (the seventh day of the week). This result can be explained by the fact that the load in this day is so weak (week-end) that we can neglect it. Indeed, this conclusion can be observed clearly from the dataset. The attribute 8 is related to temperature, the elimination of this attribute when doing feature selection mean that it hasn't a notable impact on electric load for the case of electric load in the competition studied.

B. Experiment 2:

In this experiment, the models are trained on hourly data from the NEPOOL region (courtesy ISO New England) from 2004 to 2007 (data are available on mathwork website). That is, it contains 8734 instances and 8 features as described in fig. 7. To build this experiment, we follow the same approach used in the previous example.

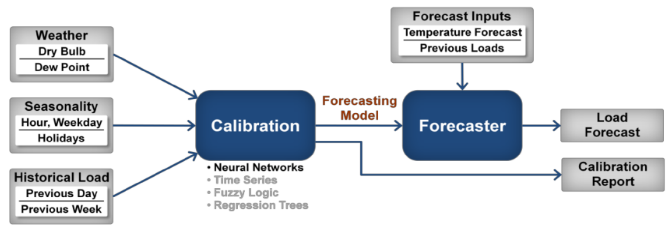

Fig. 7. Description of experiment 2

	SVR-PSO without FS	SVR-PSO with FS
MAPE	0.5296	0.0290
MSE	$1.0563.10^8$	3.4092

Table 2. Comparison of results experiment 2.

C. Discussion

We can see from the two tables of the previous section that SVR-PSO with FS have smaller MSE and MAPE than SVR-PSO without FS. That is, feature selection may improve SVR-PSO performance.

The outstanding forecasting performance of the SVR-PSO with FS technique is caused by the reason that the eliminated attributes do not have a great impact on load electricity, they can be replaced by other attributes who can have a more impact on electric load.

VII. Conclusion

In this paper, we investigate the applicability of the hybrid machine learning technique: SVR-PSO to electric load forecasting. on the one hand, we can see that the hybrid method SVR-PSO is useful for ELF. On the other hand, we can conclude that the selection of the most relevant feature can maintain the accuracy of the SVR-PSO model for forecasting. This result is useful, especially in the case of large datasets.

Future research should attempt to use more advanced methods in optimizing SVR parameters to have a better performance of the hybrid model and to determine the best way for doing feature selection.

References

[1] N. Amjady and F. Keynia, *Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm*, Energy **34** (2009), no. 1, 4901–4909.

[2] Zeyar Aung, Mohamed Toukhy, John R. Williams, Abel Sanchez, and Sergio Herrero, *Towards accurate electricity load forecasting in smart grids*, The Fourth International Conference on Advances in Databases, Knowledge, and Data Applications **DBKDA** (2012).

[3] B. E. Boser, I. M. Guyon, and V. N. Vapnik, *A training algorithm for optimal margin classiffers*, 5th Annual ACM Workshop on COLT, Pittsburgh PA (1992), 144–152.

[4] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time series analysis forecasting and control*, 3rd ed, Prentice Hall Englewood Clifs **598 pages** (1994), no. 0130607746.

[5] R. C.Eberhart and Y. Shi, *Particle swarm optimization: developments, applications and resources*, Proceedings of IEEE Congress on Evolutionary Computation **IEEE** (2001).

[6] Jinxing Che, Jianzhou Wang, and Guangfu Wang, *An adaptive fuzzy combination model based on self-organizing map and support vector regression for electric load forecasting*, Energy **37** (2012), 657–664.

[7] B.J. Chen, M.W. Chang, and C.J. Lin, *Load forecasting using support vector machines: a study on eunite competition 2001*, IEEE Trans Power Syst **19** (2004), 1821–30.

[8] Cox and Earl, *The fuzzy systems handbook a practitioners guide to building using maintaining fuzzy system*, Boston **ISBN 0-12-194270-8** (1994).

[9] Sven F. Crone and Nikolaos Kourentzes, *Feature selection for time series prediction, a combined filter and wrapper approach for neural networks*, Neurocomputing **73** (2010), 1923–1936.

[10] Pan Duan, Kaigui Xie, Tingting Guo, and Xiaogang Huang, *Short-term load forecasting for electric power systems using the pso-svr and fcm clustering techniques*, Energies **4** (2011), 173184.

[11] R.C. Eberhart and Y. Shi, *Particle swarm optimization developments applications and resources*, Proceedings of the 2001 congress on evolutionary computation (2001).

[12] Ehab Elattar, John Goulermas, and Q. Wu, *Electric load forecasting based on locally weighted support vector regression*, IEEE Transactions on Systems, Man, And Cybernetics **40** (2010), no. 4, Part C: Applications and Reviews.

[13] Ren Gang and Zhou Zhuping, *Traffic safety forecasting method by particle swarm optimization and support vector machine*, Expert Systems with Applications **38** (2011), 10420–10424.

[14] D. E. Goldberg, *Genetic algorithm in search optimization and machine learning*, Addison-Wesley **Reading** (1989).

[15] M. A. Hall, *Correlation-based feature subset selection for machine learning.*, Hamilton, New Zealand (1998).

[16] Wenwu He, Zhizhong Wang, and Hui Jiang, *Model optimizing and feature selecting for support vector regression in time series forecasting*, Neurocomputing **72** (2008), 600–611.

[17] Wei-Chiang Hong, *Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model*, Energy Conversion and Management **50** (2009), 105–117.

[18] Wei-Chiang Hong, Yucheng Dong, Chien-Yuan Lai, Li-Yueh Chen, and Shih-Yung Wei, *Svr with hybrid chaotic immune algorithm for seasonal load demand forecasting*, Energies **4** (2011), 960–977.

[19] Badar Ul Islam, *Comparison of conventional and modern load forecasting techniques based on artificial intelligence and expert systems*, IJCSI International Journal of Computer Science Issues (IJCSI) **8** (2011), no. 5, 1694–0814.

[20] Kyoung jae Kim, *Financial time series forecasting using support vector machines*, Neurocomputing **48** (1983), 311–326.

[21] Taghi Karimi, *Peak load prediction with the new proposed algorithm*, International Journal of Science and Advanced Technology **2** (2012), no. 3, ISSN 2221–8386.

[22] J. Kennedy and R.C . Eberhart, *Particle swarm optimization*, Proceedings of IEEE international conference neural networks **IEEE** (1995), 1942–8.

[23] Chi-Jie Lu, *Sales forecasting of computer products based on variable selection scheme and support vector regression*, Neurocomputing **128** (2014), 491–499.

[24] Xiaoyong Lu and Xiaomeng Geng, *Car sales volume prediction based on particle swarm optimization algorithm and support vector regression*, International Conference on Intelligent Computation Technology and Automation Shenzhen Guangdong (2011), 71–74.

[25] M. Minsky and S. Papert, *An introduction to computational geometry*, MIT Press **ISBN 0-262-63022-2** (1969).

[26] M. Mohandes, *Support vector machines for short-term electrical load forecasting*, International Journal of Energy Research **26** (2002), 335–345.

[27] A. Moraglio, C. Di Chio, and R. Poli, *Geometric particle swarm optimisation*, In Proceedings of the 10th European Conference on Genetic Programming Berlin (2007), 125–136.

[28] Minh Hoai Nguyen and Fernando de la Torre, *Optimal feature selection for support vector machines*, Pattern Recognition **43** (2010), 584–591.

[29] Dong-Xiao Niu and Ying-Chun Guo, *An improved pso for parameter determination and feature selection of svr and its application in stlf*, Multi-valued Logic (2009), 1–18.

[30] Gamse Ogcu, Omer F. Demirel, and Selim Zaim, *Forecasting electricity consumption with neural networks and support vector regression*, Procedia - Social and Behavioral Sciences **58** (2012), 1576–1585.

[31] Ping-Feng Paia and Wei-Chiang Hong, *Software reliability forecasting by support vector machines with simulated annealing algorithms*, Journal of Systems and Software **79** (2006), no. 6, 747–755.

[32] Malek Sarhani and Abellatif El Afia, *Electric load forecasting using hybrid machine learning model*, Proceeding of the 11th international conference of Intelligent Systems: Theory and Applications (Rabat, Morocco), 2014.

[33] Jamal Shahrabi, Esmaeil Hadavandi, and Shahrokh Asadi, *Developing a hybrid intelligent model for forecasting problems: Case study of tourism demand time series*, Knowledge-Based Systems **43** (2013), 112–122.

[34] S.Piramuthu, *Evaluating feature selection methods for learning in data mining applications*, European Journal of Operational Research **156** (2004), 483–494.

[35] J. W. Taylor and P. E. McSharry, *Short-term load forecasting methods: An evaluation based on european data*, IEEE Transactions on Power Systems **22** (2008), 2213–2219.

[36] Chung-Jui Tu, Li-Yeh Chuang, Jun-Yang Chang, and Cheng-Hong Yang, *Feature selection using pso-svm*, International Journal of Computer Science **33** (2007), no. 1.

[37] V.Anandhi and R.Manicka Chezian, *Forecasting the demand of pulpwood using ann and svm*, International Journal of Advanced Research in Computer Science and Software Engineering **3** (2013), no. 7, 1404–1407.

[38] V. Vapnik, S. Golowich, and A. Smola, *Support vector method for function approximation regression estimation and signal processings*, MIT Press Cambridge **9** (1992 7), 144–152.

[39] Susana M. Vieira, Lus F. Mendonca, Goncalo J.Farinha, and Joo M.C. Sousa, *Modified binary pso for feature selection using svm applied to mortality prediction of septic patients*, Applied Soft Computing **13** (2013), 3494–3504.

[40] Jianjun Wang, Li Li, and Dongxiao Niuand Zhongfu Tan, *An annual load forecasting model based on support vector regression with differential evolution algorithm*, Applied Energy **94** (2012), 65–70.

[41] Jianzhou Wang, Wenjin Zhu, Wenjin Zhu, and Donghuai Sun, *A trend fixed on firstly and seasonal adjustment model combined with the epsilon-svr for short-term forecasting of electricity demand*, Energy Policy **37** (2009), 4901–4909.

[42] L.J Wang and C. Liu, *Short-term price forecasting based on pso train bp neural network*, Electr. Power Sci. Eng. **24** (2008), 21–25.

[43] P.R. Winters, *Forecasting sales by exponentially weight moving averages*, Management Science **6** (1960), 324 – 342.

[44] Amit Kumar Yadav, Hasmat Malik, and S.S. Chandel, *Selection of most relevant input parameters using weka for artificial neural network based solar radiation prediction models*, Renewable and Sustainable Energy Reviews **31** (2014), 509–519.

[45] Wen Yu Zhang, Wei-Chiang Hong, Yucheng Dong, Gary Tsai, Jing-Tian Sung, and Guo feng Fan, *Application of svr with chaotic gasa algorithm in cyclic electric load forecasting*, Energy **45** (2012), 850–858.