Journal of Engineering and Applied Sciences 13 (17): 7440-7449, 2018
ISSN: 1816-949X

Evaluation Performance of Hybrid Localized Multi Kernel SVR (LMKSVR) in Electrical Load Data Using 4 Different Optimizations

Rezzy Eko Caraka and Sakhinah Abu Bakar
School of Mathematical Sciences, Faculty of Science and Technology,
The National University of Malaysia, Bangi, Malaysia

Abstract: The main problem using SVR is to find optimal parameter (σ) by using kernel function such as radial basis, polynomial, Gaussian and so on. Moreover, we also have to find optimal hyperplane parameter (C and ε). In the heart of statistical methods and data mining, the motivation of researcher doing this is to minimize time, money and energy in the analysis at the same time the results will be more accurate. The development of such a massive technology and the availability of data is very much making progress and improvement of methods based on data mining and machine learning. In this study, we proposed four different optimizations such as LIBSVM, MOSEK, QUADPROG, SMO applied to Localized Multi-Kernel Learning (LMKL) to assign local weights to kernel functions, so that, the best hyperplane parameters will be obtained. For the simulation, we use the electrical data and we have labeled based on the characteristics of different days (Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, National Holiday, Ramadhan). As well as we can capture the pattern of electricity consumption.

Key words: SVR, kernel, optimization, electrical, hybrid, Malaysia

INTRODUCTION

Electricity is one of the essential human needs in everyday life because almost all equipment used is a supporter of human activities using electricity as power. Today the demand for electricity is increasing significantly day. This increase is influenced by several factors such as population growth increased use of electrical equipment, technological developments increased welfare or public purchasing power and various other factors. Therefore, electric power service providers must have availability of supply and reliability of electrical energy in their distribution system to meet the electricity demand for their customers. Various studies have been conducted on electricity load forecasting. Electricity load forecasting has been a field of research for more than 50 years on operating systems and electrical planning (Srivastava *et al*., 2016). Forecasting is done to get the accuracy of electricity load forecasting for the short term, medium term and long-term. In general, electric load forecasting methods can be divided into two parts, namely: statistical methods of classical mathematics and approaches based on artificial intelligence.

Methods, researchers usually using Generalized Linear Model (Caraka *et al*. 2018; Kurniawan *et al*., 2018a, b), multiple regression, Autoregressive Moving Average Model (ARMA), Kalman Filter Technique, and Box-Jenkins to analysis load forecasting. The method has a weakness regarding ability that is not linear. With the artificial intelligence approach, such as neural network (Ko and Lee 2013), Support Vector Machine, the researchers are concerned to get higher and efficient electricity load forecasting accuracy (Kavousi-Fard and Kavousi-Fard, 2013).

Although, different in model structure and forecasting, the two categories of methods rely on historical data of electrical loads and other exogenous factors. Nowadays, it has been widely used a combination of statistical methods and artificial intelligence methods for short-term electrical load forecasting. Short-term electrical load forecasting is the prediction of load demand in the time intervals for one day to several weeks ahead (Caraka, 2016) which forecast has fundamental importance regarding plant scheduling, fuel purchase plans, security capacity, short-term maintenance as well as for the use of short-term storage. Therefore, forecasting electrical loads with high accuracy is very important for companies providing electricity services to meet the needs of electricity demand customers (Caraka *et al*., 2017a, b). Excessive predictions or forecasting of electrical loads will result in unnecessary use of power plants increased operating costs. As well as forecasting of electrical loads

Corresponding Author: Rezzy Eko Caraka, School of Mathematical Sciences, Faculty of Science and Technology, The National University of Malaysia, Bangi, Malaysia

that do not meet the electrical load requirements may cause system reliability due to insufficient power resources to meet security or cost requirements become higher due to the purchase of expensive electricity to meet the demand for electricity (Hong, 2010).

This study will utilize hybrid methods to clarify nonlinear relationships between input and output variables in short-term forecasting loads. The input variables that used in this study by using a historical data electrical load that has been clustered per day. The use of hybrid methods by utilizing the historical data of electrical loads acquired by electric service providers to obtain a more accurate short-term forecasting model, so that, the fundamental interests in plant scheduling, fuel purchase plans, security capacity, maintenance, electrical storage following the planning.

MATERIALS AND METHODS

Optimization SVR

LIBSVM: LIBSVM is an algorithm used in SVM techniques. Its application is in the classification, regression and distribution estimates. LIBSVM is also, known as a library of Support Vector Machines (SVMs). (Chang and Lin, 2011). Actively developing this package, since, the year 2000. The goal is to help users to apply SVM to their applications quickly. As we know that, the linear conditions are challenging to apply to dynamic data, therefore, need non-linear techniques to perform the analysis. Suppose there is data point (x_i, d_i) fololwing the equation (Qi, 2013):

$$d_i\left(w^T x_i + b\right) \geq +1 \text{ with the provision of } i = 1, 2, ..., N$$

The researcher did the set of non-negative slack variables $\{\xi_i\}_{i=1}^{N}$ to hyperplane, then:

$$d_i\left(w^T x_i + b\right) \geq +1\ \xi_i \text{ with the provision of } i = 1, 2, ..., N$$

MOSEK: In the field of optimization, MOSEK is a software package used to solve conic and convex. MOSEK can also be used to solve primal and dual simplex for Linear problems (LO), Quadratic (QCQO), Conic Quadratic (CQO), Semidefinite (SDO), Separable Convex (SCOPT) (Bai *et al*., 2014) and General Convex. MOSEK is also, capable of solving smooth convex nonlinear optimization of the form. Moreover, MOSEK, also, capable of solving smooth convex nonlinear optimization of the form:

$$\text{minimize } f(x) + c^T x$$
$$\text{subject to } g(x) + Ax - x^c = 0$$
$$l^c \leq x^c \leq u^c,\ l^x \leq x \leq u^x$$

where, m can be defined as constrains also n the number of variables. $x \in \mathbb{R}^n$ is a vector of decision variables, $x^{c \in \mathbb{R}m}$ is a vector of constraints or slack variables, $A \in \mathbb{R}^{m \times n}$ is the constraint matrix, $l^{c \in \mathbb{R}m}$ is them lower limit on the activity for the constraints, $u^{c \in \mathbb{R}n}$ is the upper limit on the activity for the constraints, $l^{x \in \mathbb{R}n}$ is the lower limit on the activity for the variables, $u^{x \in \mathbb{R}n}$ is the upper limit on the activity for the variables, f: $\mathbb{R}^n \to \mathbb{R}$is a non-linear function, g: $\mathbb{R}^n \to \mathbb{R}^m$ is a nonlinear vector function. In other words, it means that the ith constraint has the form:

$$l_i^c \leq g_i(x) + \sum_{j=1}^{n} a_{ij} x_j \leq u_i^c$$

when the x_i^c variable has been eliminated.

Quadratic Programming (QUADRPROG): Quadratic Programming (QUADPROG) is one of the particular programs in mathematics also, statistics. QUADPROF can be formulated as (Jorge and Wright (2006):

$$f(x) = \frac{1}{2} x' Q x + c' x$$

Assume x exists in space $\mathbb{R}^n$, Q is symmetric matrix $n \times n$ and c is any $n \times 1$ vector:

- Minimize against x
- With the limits $Ax \leq b$ and $Ex = d$

If Q is 0, then the problem will turn into linear programming. From the optimization theory, the requirement for x to be global minimizer is x must satisfy the condition of Karush-Kuhn-Tucker (KKT) (Yasin *et al*., 2016; Caraka *et al*., 2017a, b). Convex quadratic programming is a particular case of convex optimization. This method consists of two stages, first intended to convert into a convex quadratic function. Where to change the function of this goal produces a new formula that has tractable continuous relaxation with can be placed on branch-and-bound logarithms:

$$\sum_{t=1}^{T} \sum_{p=1}^{P} q_{tp} x_{tp} + \sum_{t=1}^{T-1} \sum_{t'=t+1}^{T} \sum_{p=1}^{P} \sum_{p'=1}^{P} C_{tpt'p'} x_{tpt'p'}$$

$$\sum_{p=1}^{P} X_{tpt'p'} = X_{tpt} \quad t = 1, ..., T;\ t' = 1, ..., T;$$

$$p = 1, ..., P \quad X_{tptp} = X_{tp} \quad t = 1, ..., T \quad p = 1, ..., P$$

X is a symmetric matrix (T×P) and can be considered as part of the relaxation of semidefinite programming. Any optimal solution of the dual problem will result in a convex formulation of the objective function. We can assume that (α*, u*) is the optimal solution for dual. We can produce a 0-1 quadratic formulation as follows:

$$F^{\alpha^* u^*}(x) = F(x) + \sum_{t'=1}^{T}\sum_{t'=1}^{T}\sum_{p=1}^{P}\alpha^*_{tpt'p'x\,tp}\left(\sum_{t=1}^{P} x_{t'p'}\text{-}1\right) + \sum_{t=1}^{T}\sum_{p=1}^{P} u^*_{tp}\left(x^2_{tp}\text{-}x_{tp}\right)$$

With limit:

$$\sum_{p=1}^{P} x_{tp} = 1t = 1, ..., Tx \in \{0, 1\}^{T\times P}$$

Sequential Minimal Optimization (SMO): One method that can be used to optimize hyperplane that is to solve quadratic programming with defined constraints is Sequential Minimal Optimization (SMO) (Chu and Keerthi, 2007). Sequential Minimal Optimization (SMO) is an algorithm to solve Quadratic Programming (QP) problems that arise during training in the support vector machine. SMO is a simple algorithm that can solve QP problems in SVM quickly. According to the SMO algorithm solves the QP problem in SVM without using numerical QP optimization steps. Instead, SMO chooses to resolve the smallest possible optimization problem by involving two α_i elements due to the necessity to satisfy a linear divider equation. Each step, SMO selects two α_i to be optimized together and finds the optimal values for the α_i values analytically, thereby avoiding numerical QP optimization and updating SVM to provide the latest optimum values.

Stages of hyperplane optimization in SVM include complete two Lagrange multipliers and counting b. Stages of optimization of SVM hyperplane can be described as follows.

Complete two Lagrange multipliers: QP problem in SVM training is stated as follows:

$$\max_{\alpha} W(\alpha) = \sum_{i=1}^{1}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{1} y_i y_j K(x_i x_j)\alpha_i\alpha_j$$

$$0 \le \alpha_i \le C \ \forall i \ \sum_{i=1}^{1}\alpha_i y_i = 0$$

The QP problem is solved with the SMO algorithm. A point on the equation is the optimum point when the Karush-Kuhn-Tucker (KKT) condition is satisfied and y_i, y_j $K(x_i\, x_j)$ is positive semi-definite. QP problems are met if all i:

$$\alpha_i = 0 \rightarrow y_i f(x_i) \ge 1$$
$$0 < \alpha_i < C \rightarrow y_i f(x_i) = 1$$
$$\alpha_i = C \rightarrow y_i f(x_i) \le 1$$

with $f(x_i)$ is the predicted value of x_i Data. The SMO algorithm will be lit up until the condition is met. For QP the condition of HCFs is a sufficient condition as well as a necessary condition. To complete the two Lagrange multipliers, SMO first calculates the delimiter on the multipliers. If the target y_i is not the same as the y_j target, then the following limits apply to α_j:

$$L = \max\{0, \alpha_j\text{-}\alpha_i\},\ H = \min\{C, C+\alpha_j\text{-}\alpha_i\}$$

If the target y_i is the same as the y_j target, then the following limits apply to α_j:

$$L = \max\{0, \alpha_j+\alpha_i\text{-}C\},\ H = \min\{C, \alpha_j+\alpha_i\}$$

We will look for the value of α_j To maximize the objective function W(α). If the value falls outside the boundary L and H, the value of α_j will be cut to fit within that range. Optimal α_j can be formulated as follows:

$$\alpha_j^{new} = \alpha_j + \frac{y_i(E_i\text{-}E_j)}{\eta}$$

With:

$E_i = f(x_i)\text{-}y_i$ alternatively, error training to-i

$E_j = f(x_j)\text{-}y_j$ alternatively, error training to-j

$$\eta = K(x_i, x_j) + K(x_j, x_j) \text{-} 2K(x_i, x_j)$$

Under normal conditions, η will be worth <0. In the next step, value α_j clip to within range [L, H]:

$$\alpha_j^{new,\ clipped} = \begin{cases} H, & \text{if } \alpha_j^{new} \ge H \\ \alpha_j^{new}, & \text{if } L < \alpha_j^{new} < H \\ L, & \text{if } \alpha_j^{new} \le L \end{cases}$$

After finding the solution for α_j searchable value of α_i as follows:

$$\alpha_i^{new} = \alpha_i + s\left(\alpha_j \text{-} \alpha_j^{new,\ clipped}\right)$$

with $s = y_i, y_j$

According to beyond normal conditions, η will not be negative. η worth of zero can occur if more than one training instance has the same input as vector x. On Some

Occasions, SMO will keep working even when η is not negative in this case, the objective function Ψ which must be evaluated at each end of the line segment:

$$f_i = y_i(E_i+b)-\alpha_i K(x_i, x_i)-s\alpha_j K(x_i, x_j)$$
$$f_j = y_j(E_j+b)-\alpha_i K(x_i, x_j)-\alpha_j K(x_j, x_j)$$
$$L_i = \alpha_i+s(\alpha_j-L)$$
$$H_i = \alpha_i+s(\alpha_j-H)$$
$$\Psi_L = L_i f_i+Lf_j+\frac{1}{2}L_i^2 K(x_i, x_i)+\frac{1}{2}L^2 K(x_j, x_j)+ sLL_i K(x_i, x_j)$$
$$\Psi_H = H_i f_i+Hf_j+\frac{1}{2}H_i^2 K(x_i, x_i)+\frac{1}{2}H^2 K(x_j, x_j)+ sHH_i K(x_i, x_j)$$

Calculate b: The value of b is always recalculated at each end of the stage, so that, the conditions of KKT are met for both optimization problems. The following b_i will be valid if the new α_i is not on the limit:

$$b_i = E_i+y_i(\alpha_i^{new}-\alpha_i)K(x_i, x_i)+y_i (\alpha_j^{new,\ clipped}-\alpha_j)K(x_i, x_j)+b$$

b_j will be valid if new α_j not on the limit:

$$b_j = E_j+y_i(\alpha_i^{new}-\alpha_i)K(x_i, x_j)+y_j (\alpha_j^{new,\ clipped}-\alpha_j)K(x_j, x_j)+b$$

When both b_i and b_j are valid, they are the same. If the two new Lagrange multipliers are at the limit and if L is not equal to H then the interval between b_i and b_i, all is a threshold consistent with the terms of the KKT. SMO selects the middle value between b_i and b_j

RESULTS AND DISCUSSION

Application of Localized Multiple Kernel (LMKL): Kernel machines learn a decision function regarding kernel values between training instances $\{x_i\}_{i=1}^N$ dan test instance x as follows:

$$f(x) = \sum_{i=1}^{N}\alpha_i k(x_i, x)+b$$

where researchers can use various types of kernels such as linear, polynomial, radial basis (Caraka *et al*., 2017a, b; Gonen and Alpaydin, 2011). In fact, researchers can use multi-kernel by converting the equations into:

$$k_\eta(x_i, x_j) = f_\eta\left(\{k_m(x_i^m, x_j^m)\}_{m=1}^P\right)$$

According to this, we can use combination function of f_η with linear or nonlinear kernel. Feature representations with kernel function is $\{k_m(.,.)\}(m-1)^{!p}$. P is a representative feature of the data. Where, $x_i = \{x_i\}(m = 1)^P$, $x_i^m \in \mathbb{R}^{Dm}$. Also, D_m is a dimension that describes the feature. Based on the above equation, we can get an example for a combination of more than two kernels:

$$k_\eta(x_i, x_j) = \sum_{m=1}^{P} km(x_i^m, x_j^m)$$

$$k_\eta(x_i, x_j) = \prod_{m=1}^{P} k_m(x_i^m, x_j^m)$$

The equation is also mentioned as pairwise kernel, so, it is alsol, obtained:

$$k(\{x_i^a, x_j^a\}, \{x_i^b, x_j^b\}) = k(x_i^a, x_i^b), k(x_j^a, x_j^b)+ k(x_i^a, x_j^b), k(x_j^a, x_i^b)$$

or can be written:

$$k_\eta(\{x_i^a, x_j^a\}, \{x_i^b, x_j^b\}) = \sum_{m=1}^{P} km(\{x_i^a, x_j^a\}, \{x_i^b, x_j^b\})$$

$$k_\eta(\{x_i^a, x_j^a\}, \{x_i^b, x_j^b\}) = k_\eta(\{x_i^a, x_i^b\}, k_\eta\{x_j^a, x_j^b\})+ k_\eta(\{x_i^a, x_j^b\} k_\eta\{x_j^a, x_i^b\})$$

Besides using $f_\eta(.)$ we can also use the function for parameterization by using parameter ⊖. The simplest way to parameterize is to use weights in the kernel function $\{\eta m\}^P_{m=1}$. Kernel matrix from $\omega(K)$ can be explained as to complete canonical SVM optimization using soft margin and l_1 is a slack varibales. So that:

$$\text{maximize } \omega(K) = \sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j k(x_i, x_j)$$

$$\text{concerning } \alpha \in \mathbb{R}_+^N$$

$$\text{subject to; } \sum_{i=1}^{N} \alpha_i y_j = 0$$

$$C \geq \alpha_i \geq 0 \forall_i$$

The combinations of the kernel matrix can be exposed as:

$$K_L = \left\{ K: K = \sum_{m=1}^{P} \eta_m K_m,\ K \succeq 0,\ tr(K) \leq c \right\}$$

The resulting optimization problem that minimizes the implausibility of the combined kernel matrix (the objective function value of the corresponding soft margin SVM:

$$\text{Minimize: } \omega\left(K_\eta^{tra}\right)$$

$$\text{Concerning: } K_\eta \in K_L$$

$$\text{Subject to: } tr\left(K_\eta\right) = c$$

where, K^{tra}_η is a kernel matric that is calculated only if the training data on the problem of researchers follow Semidefinite Programming (SDP) with the following equation:

$$\text{Minimize: } t$$

$$\text{Concerining: } \eta \in R^P,\ t \in R,\ \lambda \in R,\ v \in R_+^N,\ \delta \in R_+^N$$

$$\text{Subject to: } tr\left(K_\eta\right) = c$$

The main problem since the width of ϵ-margin holds the abality to affect the complexity and the generalization of the regression function indirectly. Also, It is inportant to seek another optimal value for different applications.

Application of gui matlab hybrid LMK-SVR: The process of extracting information from electrical data is divided into several stages can be seen in Fig. 1. Data cleaning, the step of filling the missing value, identifying or removing outliers and resolve inconsistent data. This stage is essential because the "NA" data can produce incorrect or unreliable outputs. Data integration, the stage of user wants to combine data from multiple sources. The data merging process is not easy because in the middle of creating data integration can be new problems such as the presence of attributes have different names on the database causing inconsistencies and redundancy

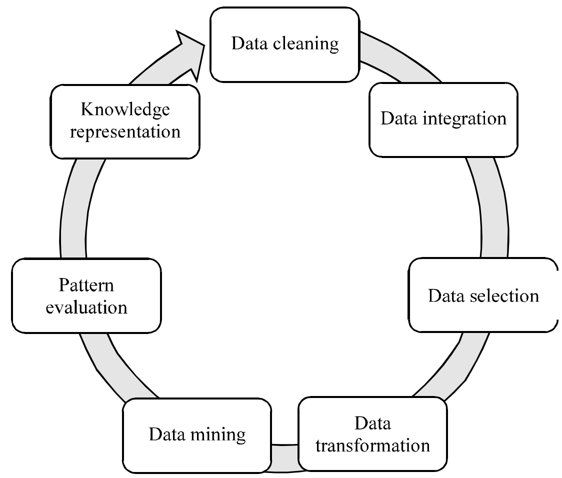

Fig. 1: Knowledge management in electrical load forecasting

of data. Data selection, the stage of the process of selecting suitable data to the process of knowledge discovery to be performed. Data transformation, the stage of the process of changing the data to fit its characteristics with the data required for the process of extracting data. Researchers can use appropriate transformations based on statistical methods usually using box-cox transformation. Data mining, the stage of data extraction process to find a pattern of the data. Pattern evaluation, the stage of the evaluation process of patterns that represent knowledge. Knowledge representation, the stage to display the knowledge that has been produced in various visual forms. The visual form aims to make the knowledge representation easy to understand.

In this simulation, we have the electrical load data which is divided between Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday, Holy Ramadhan and National Holiday. Moreover, This data cover P3BS areas UPT Banda Aceh, UPT Medan, UPT Pematang Siantar, UPT Padang, UPT Palembang, UPT Tanjung Karang, UPT North Sumatra and UPT South Sumatra. Based on Fig. 2 is a two-dimensional plot of electrical loads of the Northern Sumatera (SBU) system from 2012-2015. Also, there is a seasonal pattern especially in a national holiday with is a pattern that shows repeated changes periodically in the annual series. This plot provides information that increases the electrical load each year indicates that the electricity consumption level of customers increases for each year. Moreover, we are successful training the data by using four different optimizations such as LIBSVM, MOSEK, QUADPROG, SMO applied to Localized Kernel Learning (LKL).

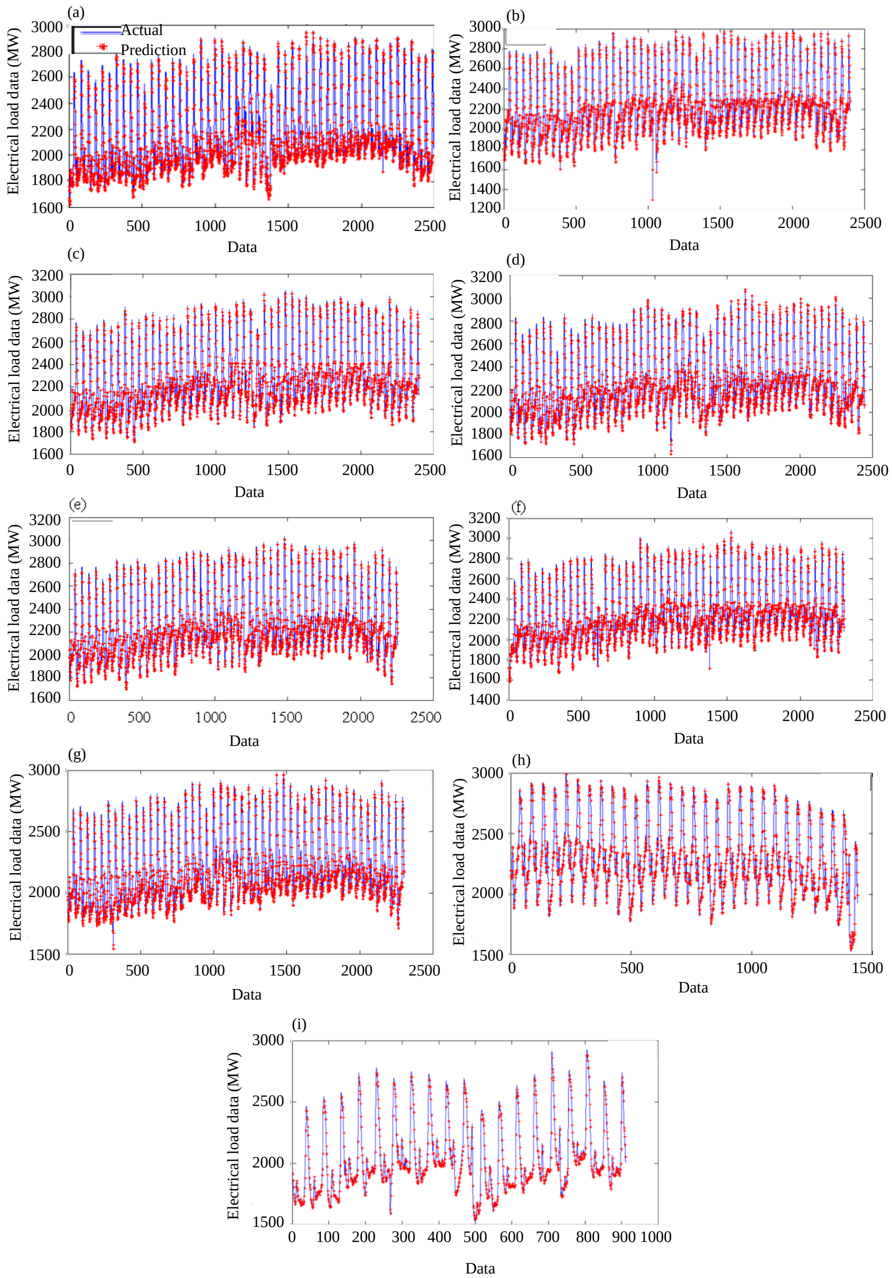

Fig. 2: Best Model LMKL-SRV with different optimization: a) Actual vs. prediction (Sunday); b) Actual vs. prediction (Monday); c) Actual vs. prediction (Tuesday); d) Actual vs. prediction (Wednesday); e) Actual vs. prediction (Thursday); f) Actual vs. prediction (Firday); g) Actual vs. prediction (Saturday); h) Actual vs. prediction (Holy Ramadhan) and i) Actual vs. prediction (National Holiday)

The pattern of data training and data testing can be seen in Fig. 2. We can justify that the data pattern has followed the same pattern, so that, the obtained LMKSVR model is feasible to use. The significance of this research, we are also, enhanced and combine GUI MATLAB Fig. 2 and 3 which is developed previous research by Yasin *et al*. (2016).

Besides, the significance of this research is that we do a combination of several types of optimization by looking at the MAPE and RMSE values as the accuracy of the model. RMSE $\sqrt{(f-o)^2}$ with f = forecast (expected values or unknown results) and o = observed values (known results):

Table 1: Optimization using LMKSVR

Data	N	Regularization	MAPE	RMSE
Monday	2400	C = 100; Tube width (e) = 0.5 Optimation = SMO, Locality = Linear	0.0249*	88.6135*
		C = 25; Tube width (e) = 0.5 Optimation = MOSEK, Locality = Quadratic	0.0249	89.1185
		C = 50; Tube width (e) = 0.5 Optimation = SMO, Locality = Quadratic	0.0250	88.6291
Tuesday	2400	C = 100; Tube width (e) = 0.5 Optimation = SMO, Locality = Linear	0.0233*	84.1870*
		C = 25; Tube width (e) = 0.5 Optimation = MOSEK, Locality = Quadratic	0.0233	84.4312
		C = 50; Tube width (e) = 0.5 Optimation = SMO, Locality = Quadratic	0.0240	82.5484
Wednesday	2400	C = 100; Tube width (e) = 0.5 Optimation = SMO, Locality = Linear	0.0236	85.0443
		C = 25; Tube width (e) = 0.5 Optimation = MOSEK, Locality = Quadratic	0.0244	82.3021
		C = 50; Tube width (e) = 0.5 Optimation = SMO, Locality = Quadrati	0.0236*	84.9959*
Thrusday	2256	C = 100; Tube width (e) = 0.5 Optimation = SMO, Locality = Linear	0.0239	84.9062
		C = 25; Tube width (e) = 0.5 Optimation= MOSEK, Locality =Quadratic	0.0241	84.1622
		C = 50; Tube width (e) = 0.5 Optimation = SMO, Locality = Quadratic	0.0239*	83.1765*
Friday	2304	C = 100; Tube width (e) = 0.5 Optimation = LIBSVM, Locality = Linear	0.0240*	86.1961*
		C = 25; Tube width (e) = 0.5 Optimation= MOSEK, Locality =Quadratic	0.0247	84.3274
		C = 50; Tube width (e) = 0.5 Optimation = SMO, Locality = Quadratic	0.0241	85.6633
Saturday	2304	C = 100; Tube width (e) = 0.5 Optimation = LIBSVM, Locality = Linear	0.0232	85.4363
		C = 25; Tube width (e) = 0.5 Optimation = MOSEK, Locality = Quadratic	0.0232*	85.3757*
		C = 50; Tube width (e) = 0.5 Optimation = SMO, Locality = Quadratic	0.0232	85.4044
Sunday	2496	C = 100; Tube width (e) = 0.5 Optimation = SMO, Locality = Linear	0.0220	85.1964
		C = 50; Tube width (e) = 0.5 Optimation = Quadrpog, Locality = Quadratic	0.0220	85.0730
		C = 25; Tube width (e) = 0.5 Optimation= LIBSVM, Locality =Linear	0.0219*	84.9332*
Holy Ramadhan	1440	C = 100; Tube width (e) = 0.5 Optimation = LIBSVM, Locality = Linear	0.0231	76.3535
		C = 50; Tube width (e) = 0.5 Optimation = Quadrpog, Locality = Quadratic	0.0230*	76.4327*
		C = 25; Tube width (e) = 0.5 Optimation = LIBSVM, Locality = Linear	0.0232	75.4842
National Holiday	912	C = 100; Tube width (e) = 0.5 Optimation = SMO, Locality = Linear	0.0229	84.6098
		C = 50; Tube width (e) = 0.5 Optimation = QUADPROG, Locality = Linear	0.0228*	84.8262*
		C = 25; Tube width (e) = 0.5 Optimation = SMO, Locality = Linear	0.0236	83.0314

*Best model

$$RMSE_{fo} = \left[\sum_{i=1}^{N} \frac{\left(Z_{fi} - Z_{oi}\right)^2}{N}\right]^{\frac{1}{2}}$$

Where:

Σ = Summation
$(z_{fi}-z_{oi})$ sup>2 = Differences, squared
N = Sample size

After the simulation results, we found that the low MAPE value can be justified the model is good to use. For more details can be in Table 1. We can classify the electrical data in Northern Sumatra as follows.

For business days: Around 04.00 am there was an increase in electrical load with a peak load at 06.30 am and this is because customers start to use electricity for prayer, cooking breakfast and preparing for work. After 6:30 am, the electricity load gradually decreased until 08.00 am. It happens because at that time the sun has risen and the lights have been extinguished.

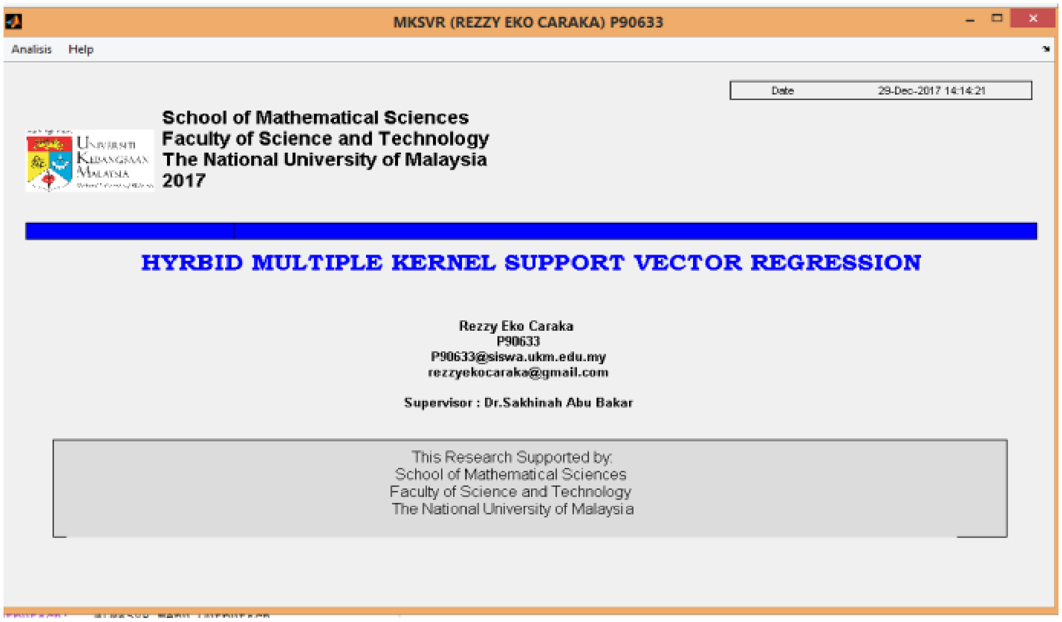

Fig. 3: Main GUI MKSVR

After 08.00 am, the load of electricity starts to increase and this happens because the community began to move, office activities and industry began to operate. During the day between 12.00 and 13.00 noon, the electrical load has decreased slightly because at this time it is the worker's break for prayer and lunch.

In the afternoon between 1800 h the load has increased dramatically, this happens because people who are at home start to turn on electrical appliances such as lights for lighting, television for entertainment and other electrical equipment used simultaneously. The peak load occurs at this time, around 7:30 pm. Starting at 21.30 people start a lot that reduces its activity, turn off the TV, reduce lights and sleep.

Saturday: On Saturday, the hourly electrical charge chart is the same as the electricity load chart on weekdays but the consumed power load is lower than the electrical load on weekdays. This happens because of a company that does not work on Saturday.

Monday, National Holiday: Expenses on Sundays differ from 08.00 am to 17.00 pm where the electricity load is very low compared to weekdays and Saturdays. The pattern of the daytime ordinary holiday loads and special holidays for Northern Sumatra from 2012 s.d. 2015 on Fig. 1 provides information that special holidays such as Ramadhan, New Year, etc. At the same hour, the electrical load is lower than the usual holiday day (Saturday and Sunday). This is due to the particular holiday industry activities and the company does not exist.

Holy Ramadhan: In the month of Holy Ramadhan electrical load has increased compared with regular months. For example, 03:00-04:30 this is because the people, especially, the Muslims have been waking up to having food and perform dawn and dawn prayers. Around 05:30 to 7:30 the electrical load in the fasting month is lower than the normal month due to the early morning activities such as preparing breakfast using electrical appliances to switch to the dawn activities.

What is interesting about the method of support vector regression is that researchers can study in the global learning and local learning. The parameters that play a significant role in building the SVR Model are the kernel. As has been explained earlier sections that the kernel that can be used very much like Gaussian, radial basis, polynomial and others. So, is the type of optimization. In the simulation using electrical data, we found no significant difference between the use of optimization. On the other hand, local learning methods do not estimate a distribution of data. Instead, they focus on extracting only the local information which is directly related to the learning task LKSVR is demonstrated to provide an automatic scheme also, systematic to locally and flexibly adapt the margin based on SVR. Besides, LKSVR may be tolerated of the noise. The significance of this research is we managed to create a MATLAB Graphical User Interface (GUI) on a Localized Multi-Kernel (LMK) can be seen in Fig. 3 and 4. We successfully built the graphic user interface for LMKSVR. In this analysis can be chosen regularization, tube width, four different

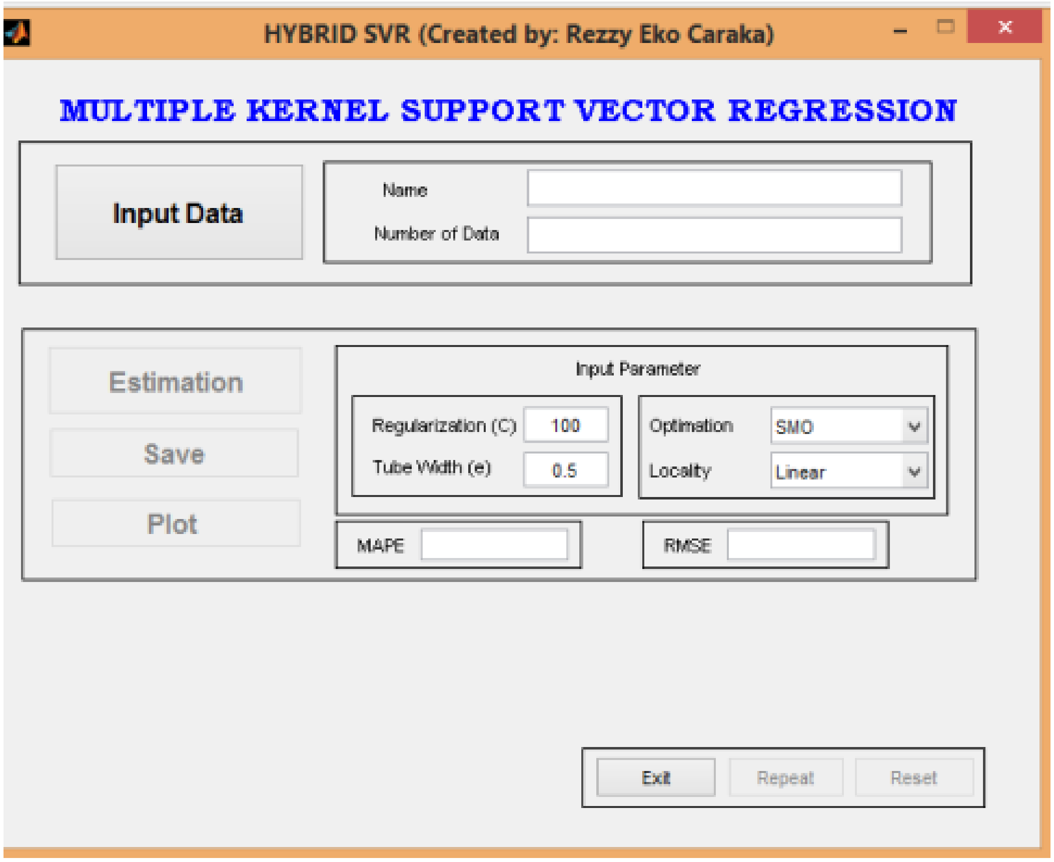

Fig. 4: GUI MKSVR

ypes of optimization and also, locality kernel. To measure performance can be justified by MAPE and RMSE.

MATLAB (Matrix Laboratory) is programming software that works with matrix concepts. MATLAB is a high-level programming language devoted to technical computing. MATLAB is developed as a programming language as well as a visualization tool that offers many capabilities to resolve cases directly related to the disciplines of mathematical science such as engineering, physics, statistics, computing and modeling. Other MATLAB capabilities that can create a GUI (Graphical User Interface) is a Window-based application.

The GUI is a graphical display medium instead of a text command (syntax) for users interacting, so, the GUI is called Window-based applications. Creating window-based applications with MATLAB can be done using the GUI Designer (GUIDE). The GUI Designer provides a place to design the GUI as well as the components required in making the GUI.

Figure 3 is the primary interface before doing the analysis. It can be seen that in this interface there is the analysis and help button. If the researcher pressing the analysis button will appear a new interface that can be seen in Fig. 4. On this interface, researchers can input data as well as choose the parameter regularization (c), tube width (e), optimization and locality kernel to modeling based on localized multiple kernel support vector regression.

CONCLUSION

Based on the results achieved, it can be concluded that MKSVR Model is excellent at forming electrical load model. It can be seen from the value of MAPE generated all of them <2%. Local learning adopts a mostly different way to construct parameters. This type of learning is even more task-oriented than minimum error minimax probability machine and maximal conditional learning.

ACKNOWLEDGEMENT

This research was supported by School of Mathematical Sciences, The National University of Malaysia GUP-2017-118.

REFERENCES

Bai, L., J.E. Mitchell and J.S. Pang, 2014. Using quadratic convex reformulation to tighten the convex relaxation of a quadratic program with complementarity constraints. Optim. Lett., 8: 811-822.

Caraka, R.E., 2016. Ramadhan short-term electric load: A hybrid model of Cycle Spinning Wavelet and Group Method Data Handling (CSW-GMDH). Proceedings of the 2nd International Conference on Actuarial Science and Statistics (ICASS), December 5-6, 2016, Statistics Diponegoro University, Semarang, Indonesia, pp: 1-16.

Caraka, R.E., H. Yasin and A.W. Basyiruddin, 2017b. Peramalan Crude Palm Oil (CPO) menggunakan support vector regression kernel radial basis. Mat., 7: 43-57.

Caraka, R.E., S. Shohaimi, I.D. Kurniawan, R. Herliansyah and A. Budiarto *et al*., 2018. Ecological show cave and wild cave: Negative binomial Gllvm's arthropod community modelling. Procedia Comput. Sci., 135: 377-384.

Caraka, R.E., S.A. Bakar, B. Pardamean and A. Budiarto, 2017a. Hybrid support vector regression in electric load during national holiday season. Proceedings of the 2017 International Conference on Innovative and Creative Information Technology (ICITech'17), November 2-4, 2017, IEEE, Salatiga, Indonesia, ISBN:978-1-5386-4047-0, pp: 1-6.

Chang, C. and C. Lin, 2011. LIBSVM: A library for support vector machines. ACM. Trans. Intell. Syst. Technol., 2: 1-39.

Chu, W. and S.S. Keerthi, 2007. Support vector ordinal regression. Neural Comput., 19: 792-815.

Gonen, M. and E. Alpaydin, 2011. Multiple kernel learning algorithms. J. Mach. Learn. Res., 12: 2211-2268.

Hong, T., 2010. Short term electric load forecasting. Ph.D Thesis, North Carolina State University, Raleigh, North Carolina.

Jorge, N. and S.J. Wright, 2006. Quadratic Programming. In: Numerical Optimization, Jorge, N. and S. Wright (Eds.). Springer, New York, NY., USA., ISBN:978-0-387-30303-1, pp: 448-492.

Kavousi-Fard, A. and F. Kavousi-Fard, 2013. A new hybrid correction method for short-term load forecasting based on ARIMA, SVR and CSA. J. Exp. Theor. Artif. Intell., 25: 559-574.

Ko, C.N. and C.M. Lee, 2013. Short-term load forecasting using SVR (support vector regression)-based radial basis function neural network with dual extended Kalman filter. Energy, 49: 413-422.

Kurniawan, I.D., C. Rahmadi, R.E. Caraka and T.E. Ardi, 2018b. Cave-dwelling Arthropod community of Semedi Show Cave in Gunungsewu Karst Area, Pacitan, East Java, Indonesia. Biodiversitas, 19: 857-866.

Kurniawan, I.D., R.H. Soesilohadi, C. Rahmadi, R.E. Caraka and B. Pardamean, 2018a. The difference on Arthropod communities' structure within show caves and wild caves in Gunung Sewu Karst Area, Indonesia. Ecol. Environ. Conserv., 24: 81-90.

Lanckriet, G.R.G., N. Cristianini, P. Bartlett, L.E. Ghaoui and M.I. Jordan, 2004. Learning the kernel matrix with semide?nte programming. J. Mach.Learn. Res., 5: 27-72.

Qi, S., 2013. Support vector machine and its realization through LIBSVM: A superficial review. J. Mach. Learn., Vol. 1,

Srivastava, A.K., A.S. Pandey and D. Singh, 2016. Short-term load forecasting methods: A review. Proceedings of the International Conference on Emerging Trends in Electrical Electronics & Sustainable Energy Systems (ICETEESES'16), March 11-12, 2016, IEEE, Sultanpur, India, ISBN:978-1-5090-2119-2, pp: 130-138.

Yasin, H., R.E. Caraka and A. Hoyyi, 2016. Prediction of crude oil prices using Support Vector Regression (SVR) with grid search-cross validation algorithm. Global J. Pure Appl. Math., 12: 3009-3020.