PRICE PREDICTIVE ANALYSIS MECHANISM UTILIZING GREY WOLF OPTIMIZER-LEAST SQUARES SUPPORT VECTOR MACHINES

Zuriani Mustaffa[1] and Mohd Herwan Sulaiman[2]
[1]Faculty of Computer Systems & Software Engineering, Universiti Malaysia Pahang, Lebuhraya Tun Razak, Gambang, Kuantan, Pahang, Malaysia
[2]Sustainable Energy & Power Electronics Research Cluster, Fakulti Kejuruteraan Elektrik & Elektronik, Universiti Malaysia Pahang, Pekan, Pahang, Malaysia
E-Mail: zuriani@ump.edu.my

ABSTRACT
A good selection of Least Squares Support Vector Machines (LSSVM) hyper-parameters' value is crucial in order to obtain a promising generalization on the unseen data. Any inappropriate value set to the hyper parameters would directly demote the prediction performance of LSSVM. In this regard, this study proposes a hybridization of LSSVM with a new Swarm Intelligence (SI) algorithm namely, Grey Wolf Optimizer (GWO). With such hybridization, the hyper-parameters of interest are automatically optimized by the GWO. The performance of GWO-LSSVM is realized in predictive analysis of gold price and measured based on two indices viz. Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSPE). Findings of the study suggested that the GWO-LSSVM possess lower prediction error rate as compared to three comparable algorithms which includes hybridization models of LSSVM and Evolutionary Computation (EC) algorithms.

Keywords: gold price predictive analysis, grey wolf optimizer, least square support vector machines.

INTRODUCTION
The unpredictability of gold price in recent years has attracted much attention from many parties which includes commodities traders, mining companies, investors and academia as well. Govern by high nonlinearity features, the price of gold has experienced an expeditious increases during the last several years [1] and is continually predicted to be on steady condition in 2015 [2]. Nonetheless, to accurately predict the price of gold is such a great challenge. With the uncertainties of world economic and surrounded by various factors, this challenge has paved a positive way for academic community in exploring a new method for predictive analysis purposes.

In literature, the application of well-known machine learning algorithm, namely Artificial Neural Network (ANN) has been proposed for the said task and this includes for gold price [3, 4]. Nonetheless, the adaptation of Empirical Risk Minimization (ERM) which tends to minimize the training error rather than the true error makes the ANN suffer with over fitting problem [5]. Recently, the emergence of Statistical Learning Theory based algorithm, viz. Least Squares Support Vector Machines (LSSVM) [6] has been a rival to the ANN. As opposed to ANN, LSSVM which is a modification version of Support Vector Machines (SVM) [7] adopted Structural Risk Minimization (SRM) principle which minimizes the generalization error instead of training error [8]. Hence, promising generalization can be obtained.

Owing to its remarkable nonlinear mapping capabilities, LSSVM has been proven to contribute a significant impact in solving various data mining tasks which includes prediction, classification and many others [9, 10]. However, despite its diversity in application, it is well documented that the generalization performance of LSSVM is highly dependent on the value of its two free hyper-parameters, namely regularization parameter, γ and kernel parameter, $\sigma 2$ [11]. Any improper value set to the hyper parameters would result in undesired generalization performance.

An extensive literature reviews reveals that a good numbers of meta-heuristic algorithms have been proposed in order to cater this matter. In [11-13], the LSSVM is hybrid with a Swarm Intelligence (SI) algorithm, namely Particle Swarm Optimization (PSO) for parameter tuning. In the studies, the efficiency of PSO-LSSVM is realized in different problem domain which includes nuclear science, shipping and water drainage and irrigation respectively. On the other hand, the capability of Genetic Algorithm (GA), which is a dominant algorithm in Evolutionary Algorithm (EA) class was tested in several function estimation problems, which includes in [14, 15]. Meanwhile, in [16], the LSSVM is hybridized with Fruit Fly Optimization (FFO) [17] for electric load predictive analysis. In the study, the FFO which is inspired from the food searching behaviour is employed as an optimizer to LSSVM. Later, the FFO-LSSVM is compared against several identified techniques which include single LSSVM and regression technique. Final results suggested that the FFO-LSSVM is capable to produce lower error rate relative to several identified metrics.

With respect to that matter, in this study, the LSSVM is optimized utilizing a new SI algorithm, namely Grey Wolf Optimizer (GWO) [18] which was originally inspired by the collective behaviour of grey wolf. This algorithm has been proven to be competitive to the existing meta-heuristic algorithms which include the PSO, GSA, Differential Evolution (DE), Evolutionary Programming (EP) and Evolution Strategy (ES) [18]. With such an impressive performance, in this study, the GWO is utilized as an optimization tool for the said hyper parameters.

REGRESSION BASED ON LEAST SQUARES SUPPORT VECTOR MACHINES

Proposed by [6], LSSVM is an alteration of the original SVM. As a modified version of standard SVM, LSSVM uses square errors instead of nonnegative errors in the cost function and applies equality constraint rather inequality constraint of SVM in the problem formulation. As a result, one solves linear equations instead of Quadratic Programming (QP) solver which in practice is harder to use [6]. The adaptation of QP also raises computational complexity in training which is at least quadratic with respect to the number of training data [19]. With the reformulation of LSSVM, it simplifies complex calculation which led to easier and faster training task. Hence, a simpler optimization problem can be obtained.

Formally, given a training set of N points $\{x_i, y_i\}^N$ with the input values x_i and the output values y_i, for nonlinear regression, the data are generated by the nonlinear function $y(x) = f(x_i) + e_i$, the objective is to estimate a model of the following form [6].

$$y(x) = w^T \varphi(x_i) + b + e_i \tag{1}$$

Where w is the weight vector, $\varphi(.):R^n$ is the nonlinear function which maps the input space into a higher dimensional feature space, b denotes the biasness and e_i is the error between the actual and predicted output at the ith training data. The input, x_i and output, $y(x)$ are described in next Section. The coefficient vector w and biased term b can be obtained through the optimization problem which is formulated as follows [6]:

$$\min_{w,b,e} J(w,e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{i=1}^{N} e_i^2 \tag{2}$$

subject to the equality constraints

$$y_i = w^T \varphi(x_i) + b + e_i, \quad i = 1, 2, \ldots, N$$

The first part of (2) is used to regulate the weight sizes and penalize large weights. On the other hand, the second part of (2) indicates the error in training data.

Applying the Lagrangian multiplier to (2) yields:

$$L(w,b,e;\alpha) = J(w,e) - \sum_{i=1}^{N} \alpha_i \{w^T \varphi(x_i) + b + e_i - y_i\} \tag{3}$$

where α_i are Lagrange multipliers called support values that can be positive or negative in LSSVM formulation due to the equality constraints, γ is the regularization parameter which balances the complexity of the LSSVM model, i.e. $y(x)$, and the training error. Differentiating (3) with w, b, e_i and α_i, the conditions for optimality of this problem can be obtained by setting all derivatives equal to zero, as expressed in the following:

$$\begin{aligned}
&\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{N} \alpha_i \varphi(x_i) \\
&\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{N} \alpha_i = 0 \\
&\frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i \\
&\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow w^T \varphi(x_i) + b + e_i - y_i = 0
\end{aligned}$$

$$i = 1, 2, \ldots, N \tag{4}$$

By elimination of w and e_i, the optimization problem can be transformed into the following linear equations:

$$\begin{bmatrix} 0 & y^T \\ y & \Omega + I/\gamma \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 1_v \end{bmatrix} \tag{5}$$

With $y = [y1; \ldots; y_N]$, $\alpha=[\alpha_1, \ldots, \alpha_N]$, i is the identity matrix and $1_v = [1;\ldots;1]$. The kernel trick is applied as follows:

$$\begin{aligned} \Omega il &= \varphi(x_i)^T \varphi(x_l) \\ &= K(x_i, x_l) \qquad i, l = 1, \ldots, N \end{aligned} \tag{6}$$

The resulting of LSSVM model for regression in (1) becomes:

$$y(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i) + b \tag{7}$$

where α and bare the solutions of (5). In (7), there are several available kernel functions $K(x, x_i)$, namely Gaussian kernel or Radial Basis Function (RBF) kernel, Polynomial kernel or Linear kernel function. In this study, the RBF kernel is used since its suitability in dealing with nonlinear cases [19, 20] and gives good performance in many prediction cases [21]. It is expressed as:

$$K(x, x_i) = e^{\frac{\|x - x_i\|^2}{2\sigma^2}} \tag{8}$$

where σ^2 is a tuning parameter which is associated with RBF kernel. By using kernel function, it allows the data which are not linearly separable in input space to become linearly separable in high dimensional feature space. Another tuning parameter, which is regularization parameter, γ can be seen in (2).

OPTIMIZATION BASED ON GREY WOLF OPTIMIZER

Classified as SI algorithm, GWO [18] is a new nature-inspired meta-heuristic optimization algorithm that is motivated from intelligent behaviour of grey wolf. In nature, the grey wolf is considered as apex predators and populates at the top of the food chain. In GWO, the grey wolf population is ranked into 4 hierarchies, viz. alpha, beta delta and omega. The top level, namely the he alpha incorporated of male and female grey wolf. The duty of alpha is to make decision making such as about hunting, sleeping place and others. As a leader, the alpha is placed at the top of the hierarchy. It is worth noting that, the dominant role is measures based on managing wise, not the strength.

The second level namely beta responsible in assisting the alpha in decision making or any other activities occur in the pack. Similarly like alpha, the beta can be male or female. As a consultant to alpha, the beta would be the best suitor for replacement in alpha if one of the alpha deceased or become old. The beta also plays role as advisor for the alpha in managing discipline of the pack. Meanwhile, the delta, have to submit the solution to alpha and beta but they dominate the omega. This group consist of scouts, sentinels, elders, hunters and caretakers. Lastly, the omega, which ranked last in the hierarchy, plays the role as scapegoat.

Social hierarchy

In GWO, the fittest solution is represented by alpha (α), followed by the second and third best solutions namely beta (β) and delta (δ) respectively. Meanwhile, the rest of the candidate solutions are considered as omega (ω). The hunting which represents the optimization process is guided by α, β and δ while the ω follows the three groups.

Encircling prey

During hunting, the wolves tend to encircle their prey. As to model the encircling prey, the following equation is used:

$$\vec{D} = |\vec{C}.\vec{X}_p(t) - \vec{X}(t)| \quad (9)$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A}.\vec{D} \quad (10)$$

where t =current iteration, $\vec{A}$ and $\vec{C}$ = coefficient vectors, $\vec{X}_p$ = position vector of the prey and $\vec{X}$ = position vector of the grey wolves. For vectors $\vec{A}$ and $\vec{C}$, is calculated as follows:

$$\vec{A} = 2\vec{a}.\vec{r}_1 - \vec{a} \quad (11)$$

$$\vec{C} = 2.\vec{r}_2 \quad (12)$$

where components of $\vec{a}$ are linearly decreased from 2 to 0 over the curse of iterations. Meanwhile, r_1 and r_2 are random vectors in the range of [0, 1].

Hunting

Commonly, the hunting is guided by the alpha. However, both beta and delta might also involved in hunting occasionally. The alpha which represents the fittest candidate solution, beta and delta are the experts about the potential location of prey. Thus, the first three best solutions obtained so far are saved while the other agents (including omegas) are induced to update their positions based on the position of the best search agents. This is defined by:

$$\vec{D}_\alpha = |\vec{C}_1.\vec{X}_\alpha - \vec{X}|, \vec{D}_\beta = |\vec{C}_2.\vec{X}_\beta - \vec{X}|, \vec{D}_\delta = |\vec{C}_3.\vec{X}_\delta - \vec{X} \quad (13)$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1.(\vec{D}_\alpha), \vec{X}_2 = \vec{X}_\beta - \vec{A}_2.(\vec{D}_\beta), \vec{X}_3 = \vec{X}_\delta - \vec{A}_3.(\vec{D}_\delta) \quad (14)$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (15)$$

HYBRID GWO-LSSVM

As highlighted previously, the generalization performance of LSSVM is highly dependent on the value of its hyper-parameters, viz. γ and σ^2. To ensure the optimality of value of hyper-parameters, in this study, the LSSVM is hybridized with GWO algorithm.

In GWO-LSSVM algorithm, the GWO algorithm is utilized as an optimization tool for the LSSVM hyper-parameters. Generally, the LSSVM is integrated with the GWO algorithm, where here, the predicted results obtained by LSSVM acts as a fitness function evaluation. The optimized value of LSSVM hyper-parameters can be achieved after a maximum number of iteration has been reached. In this study, the objective function is served by Mean Absolute Percentage Error (MAPE), where the lower the MAPE, the better the prediction accuracy (see section 5).

Before the initialization, the normalized training and validation data is fed into the prediction model.

Initialization

Initially, the number of search agents (grey wolves) and the maximum iteration are set. Figure-1 shows the example of possible solutions which are the parameters of interest in X.

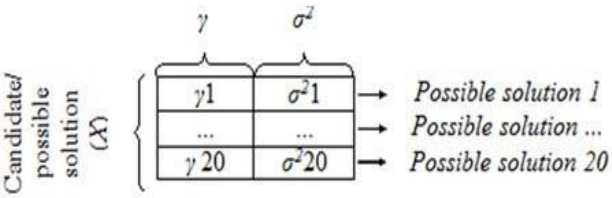

Figure-1. Representation of search agent position as possible solutions in *X*.

The number of search agent/candidate solution indicates the possible solutions which are the parameters to be optimized, namely the hyper-parameters, γ and σ^2. In this study, the number of possible solutions position is set to 20.

Evaluation

The objective function of the hyper-parameters is evaluated based on training and validation set using LSSVM function which is integrated in GWO algorithm. In this study, the objective function in guided by Mean Absolute Percentage Error (MAPE). The goal is to find the ideal combination of hyper-parameters (i.e. possible solutions) that will minimize the MAPE, as follows:

$$MAPE = \frac{1}{N}\left[\sum_{n=1}^{N}\left|\frac{y_n - y(x_n)}{y_n}\right|\right] \tag{16}$$

where $n = 1, 2, \ldots, N$
y_n= actual values
$y(x_n)$ = predicted values / approximate values by predictor models
N =Number of test data

The best result of MAPE is kept as alpha's score; the second best result is kept as beta's best score while the third best result is kept as delta's score.

METHODOLOGY

An ideal predictive analysis model should include all the possible elements/input which influences the output. Nonetheless, it is worth noting that it is impossible to include all the necessary inputs due to several constraints such as the unavailability of the data. In this study, the time series data employed are high frequency (daily) data. Thus, other economic factor such as industrial production, inflation, Gross Domestic Product (GDP), and oil supply and demand inventory are excluded as they are only available on monthly or quarterly [22, 23].

In this study, four time series data of non-renewable natural resources commodities price were employed as inputs. They are gold price (GC), silver price (SI), palladium price (PA) and crude oil price (CL) while the output was GC from day 21 onwards. These data are obtained from Barchart website [24]. The time series utilized covered from January 2009 to October 2012. The time series data were chosen due to its high correlation between them which is believed to influence each other's price [25]. The correlation between time series data is as tabulated in Table-1.

Table-1. Price correlation.

	GC	SI	PA	CL
GC	1	-	-	-
SI	0.9071	1	-	-
PA	0.8449	0.9076	1	-
CL	0.8107	0.8372	0.8642	1

In this study, the input variables are daily closing spot price, percent change in daily closing spot price from the previous day, standard deviation over the previous 5 trading days and standard deviation over the previous 21 trading days of gold, silver, palladium and crude oil. Meanwhile for the output variable is daily spot price of gold from day 21 onwards (GC21). Table-2 indicates the variables assigned to the inputs involved. On the other hand, for training, validation and testing, the data is divided into 70:15:15 respectively. The data is divided in that manner since this is the best proportion that able to give the desired result for the case under study. Table-3 shows the proportion of the samples for the said data set.

Table-2. Input and output variables.

Input	Variable	Output
Daily closing price of gold, silver, palladium and crude oil price	GC, SI, PA and CL	GC21
Percent change (%Chg) in daily closing spot prices from the previous day of GC, SI, PA and CL	%GC, %SI, %PA and % CL	
Standard deviation over the previous 5 days trading days of GC, SI, PA and CL	GC5, SI5, PA5 and CL5	
Standard deviation over the previous 21 days trading days of GC, SI, PA and CL	GC21, SI21, PA21 and CL21	

Table-3. Training, validation and testing.

	(%)	Samples (Days)
Training	70	1-661
Validation	15	662-802
Testing	15	803-943

RESULTS AND DISCUSSION

The experiments were designed to evaluate the efficiency of the hybrid GWO-LSSVM in predictive analysis the time series data of interest. The modeling process is carried out by MATLAB R2010a by using LS-SVMlab Toolbox [26].

There are two stages of experiment that have been conducted in this study. The first stage is to run the GWO-LSSVM using raw data and normalized data. The objective is to identify which data sets will produce the best result for GWO-LSSVM in prediction task. Table-4 shows the obtained result from the experiment. From the table, it clearly shows that the GWO-LSSVM capable to give better performance by using normalized data. This is indicates by lower MAPE produced in testing, which is 5.4534%. The difference between the MAPE obtained by using raw data is 13.7998%. Later, the experiment is proceeded by using the obtained results from normalized data. In this experiment, the performance of GWO-LSSVM is compared against the ones produced by hybrid LSSVM with two meta-heuristic algorithm viz. Artificial Bee Colony algorithm (ABC-LSSVM), Genetic Algorithm (GA-LSSVM). Besides the hybrid algorithm, the comparison against the results produced by single GWO is also performed. The properties of GWO, ABC and GA are as tabulated in Table-5. On the other hand, Table-6 presents the results of GC price prediction produced by all identified algorithms.

Table-4. GC price predictive analysis for GWO-LSSVM: raw data vs. normalized data.

	Raw Data	Normalized Data
γ	11.2306	1.8391
σ^2	993.299	148.5318
MAPE Training (%)	0.6786	3.1875
MAPE Validation (%)	19.3341	4.4143
MAPE Testing (%)	19.2532	5.4534
RMSPE (%)	0.1954	0.0678

Table-5. Properties of GWO, ABC and GA.

	GWO	ABC	GA
Population size	20	20	20
Maximum Iteration	100	100	100
Crossover probability	-	-	0.9
Mutation probability	-	-	0.1

Table-6. Gold price predictive analysis: GWO-LSSVM vs. ABC-LSSVM vs.GA-LSSVM vs. GWO.

	GWO-LSSVM	ABC-LSSVM	GA-LSSVM	GWO
γ	1.8391	1000	1	-
σ^2	148.5318	15.5884	133.2749	-
x1	-	-	-	0.8292
x2	-	-	-	0
x3	-	-	-	0.0272
x4	-	-	-	0.0002
x5	-	-	-	0
x6	-	-	-	0
x7	-	-	-	0.0014
x8	-	-	-	0.0005
x9	-	-	-	0
x10	-	-	-	0.0002
x11	-	-	-	0
x12	-	-	-	0.0002
x13	-	-	-	0.0006
x14	-	-	-	0
x15	-	-	-	0.0002
x16	-	-	-	0.0002
MAPE Training (%)	3.1875	0.1401	3.4116	4.7709
MAPE Validation (%)	4.4143	1.6336	4.7189	4.6428
MAPE Testing (%)	5.4534	7.1158	6.8456	2.9564
RMSPE (%)	0.0678	0.0910	0.0809	0.0367

From the table, it can be seen that the proposed GWO-LSSVM capable to produce 5.4534% of MAPE in testing. This is followed by GA-LSSVM with 6.8456% while 7.1158% was recorded by ABC-LSSVM. On the other hand, single GWO yielded smallest MAPE which is 2.9564%. However, even though good result is obtained, such an approach is inefficient due to the need of variables assignment; one parameter has to be assigned to each variable. Hence, the more variables mean the more parameters to be optimized. This will make the optimization process more complicated. In addition, if there is a change on the number of input, it might cause the related part the algorithm to change as well.

CONCLUSIONS

In this study, a hybridization of GWO-LSSVM is proposed for gold price predictive analysis. The main objective is to obtain an optimal value for the free hyper parameters of LSSVM in order to achieve good generalization. For comparison purposes, three algorithms were identified, namely ABC-LSSVM, GA-LSSVM single GWO. Based on the simulations which were adopted from real gold price, it can be said that the GWO-LSSVM capable to produce promising error rate with efficient approach as compared to the other comparison algorithms. With that, it is safe to conclude that the GWO-LSSVM is more reliable and can be used as a prediction tool for the said time series data.

REFERENCES

[1] C. Pierdzioch. 2014. The International Business Cycle and Gold Price Fluctuations. The Quarterly Review of Economics and Finance. Vol. 54, pp. 292-305.

[2] D. Carlson. 2014. Gold Prices Likely To Start 2015 Weak, But Rise Into Year End. Available: www.forbes.com/sites/kitconews/2014/12/15/gold-prices-likely-to-start-2015-weak-but-rise-into-year-end/

[3] Mohd Hussein S. F., Nor Shah M.B., Abd Jalal M.R and Abdullah S.S. 2011. Gold Price Prediction Using Radial Basis Function Neural Network. In: Proceedings of the 4th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO). pp. 1-11.

[4] V. Varahrami. 2011. Recognition of Good Prediction of Gold Price Between MLFF and GMDH Neural Network. Journal of Economics and International Finance. Vol. 3, pp. 204-210.

[5] J. Cheng, J. Qian and Y.Guo. Least Squares Support Vector Machines for Gas Concentration Forecasting in Coal Mine. International Journal of Computer Science and Network Security. Vol. 6, pp. 125-129.

[6] Suykens J. A. K., Van Gester, T., De Brabanter, J., De Moor B. and Vandewalle J. 2002. Least Squares Support Vector Machines. Leuven, Belgium: World Scientific Publishing Co. Pte. Ltd.

[7] Vapnik V. N. 1995. The Nature of Statistical Learning Theory 2nd ed. ed. New York: Springer-Verlag.

[8] K. O. Elish and M. O. Elish. 2008. Predicting defect-prone software modules using Support Vector Machines. The Journal of Systems and Software. Vol. 81, pp. 649-660.

[9] Cheng G. Guo, R. and Wu W. 2010. Petroleum Lithology Discrimination Based on PSO-LSSVM Classification Model. In: Proceedings of the Second International Conference on Computer Modeling and Simulation (ICCMS). pp. 365-368.

[10] Li Y. 2009. Short-Term Load Forecasting Based on LS-SVM Optimized by BCC Algorithm. In: Proceedings of the 15th International Conference on Intelligent System Applications to Power Systems (ISAP). pp. 1-5.

[11] B. T. Jiang and F. Y. Zhao. 2013. Particle Swarm Optimization-based Least Squares Support Vector Regression for Critical Heat Flux Prediction. Annals of Nuclear Energy. Vol. 53, pp. 69-81.

[12] Zhou B. and Shi A. 2010. LSSVM and Hybrid Particle Swarm Optimization for Ship Motion Prediction. In: Proceedings of the International Conference on Intelligent Control and Information Processing (ICICIP). pp. 183-186.

[13] Xiang Y. and Jiang L. 2009. Water Quality Prediction Using LS-SVM and Particle Swarm Optimization. In: Proceedings of the Second International Workshop on Knowledge Discovery and Data Mining (WKDD). pp. 900-904.

[14] M. H. Sulaiman, M.W. Mustafa, O. Aliman, S.N. Abd. Khalid and H. Shareef. 2012. Real and Reactive Power Flow Allocation in Deregulated Power System Utilizing Genetic-Support Vector Machines Technique. International Review of Electrical Engineering. Vol. 5, pp. 2199-2208.

[15] Z. Yang, X.S. Gu, X.Y. Liang and L.C. Ling. 2010. Genetic Algorithm-Least Squares Support Vector Regression based Predicting and Optimizing Model on Carbon Fiber Composite Integrated Conductivity. Materials and Design, Vol. 31, pp. 1042-1049.

[16] H. L. 2012. Annual Electric Load Forecasting by a Least Squares Support Vector Machines with a Fruit Fly Optimization Algorithm. Energies. pp. 4430-4445.

[17] W. T. Pan. 2012. A New Fruit Fly Optimization Algorithm: Taking the Financial Distress Model as an Example. Knowledge Based System. pp. 69-74.

[18] S. Mirjalili, S. M. Mirjalili and A. Lewis. 2014. Grey Wolf Optimizer. Advances in Engineering Software. Vol. 69, pp. 46-61.

[19] G. B. Huang, H. Zhou, Ding X. and Zhang R. 2012. Extreme Learning Machine for Regression and Multiclass Classification," IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics. Vol. 42.

[20] D. Wu, Feng S. and He Y. 2007. Infrared Spectroscopy Technique for the Nondestructive measurement of Fat Content in Milk Powder. Journal Dairy Science. Vol. 90, pp. 3613-3619.

[21] R. Liao, H. Zheng, S. Grzybowski and L. Yang. 2011. Particle Swarm Optimization-Least Squares Support Vector Regression based Forecasting model on Dissolved Gases in Oil-Filled Power Transformer. Electric Power Systems Research. Vol. 81, pp. 2074-2080.

[22] S. Hammoudeh, R. Sari and B. T. Ewing. 2008. Relationships Among Commodities and with Financial Variables: A New Look. Contemporary Economic Policy. Vol. 27, pp. 251-264.

[23] S. Kulkarni and I. Haidar. 2009. Forecasting Model for Crude Oil Price Using Artificial Neural Networks and Commodity Futures Prices. International Journal of Computer Science and Information Security. 2.

[24] Barchart. 2012 Available: http://www.barchart.com/historicalquote.php?view=quote&sym=CLY00&txtDate=12%2F31%2F09&submit=Get+Data.

[25] J. Ahmed Farid and F. Salahudin. 2010. Risk Frameworks and Applications, 2nd ed. Karachi: Alchemy Technologies Pvt. Ltd.

[26] K. Pelkmans, J.A.K. Suykens, T.V. Gestel, J. De brabanter, L. Lukas, B. Hamer, B. De Moor and J. Vandewalle. 2002. LS-SVMlab: A Matlab/C Toolbox for Least Squares Support Vector Machines. Available: www.esat.kuleuven.be/sista/lssvmlab/