

LSSVM Parameters Tuning with Enhanced Artificial Bee Colony

Zuriani Mustaffa and Yuhanis Yusof
School of Computing, University Utara Malaysia, Malaysia

Abstract: To date, exploring an efficient method for optimizing Least Squares Support Vector Machines (LSSVM) hyper-parameters has been an enthusiastic research area among academic researchers. LSSVM is a practical machine learning approach that has been broadly utilized in numerous fields. To guarantee its convincing performance, it is crucial to select an appropriate technique in order to obtain the optimized hyper-parameters of LSSVM algorithm. In this paper, an Enhanced Artificial Bee Colony (eABC) is used to obtain the ideal value of LSSVM's hyper parameters, which are regularization parameter, γ and kernel parameter, σ^2 . Later, LSSVM is used as the prediction model. The proposed model was employed in predicting financial time series data and comparison is made against the standard Artificial Bee Colony (ABC) and Cross Validation (CV) technique. The simulation results assured the accuracy of parameter selection, thus proved the validity in improving the prediction accuracy with acceptable computational time.

Keywords: ABC, LSSVM, financial time series prediction, parameter tuning.

Received November 1, 2011; accepted May 22, 2012; published online April 4, 2013

1. Introduction

Recently, a new and promising data mining technique, Least Square Support Vector Machines (LSSVM), which has been proposed by Suykens *et al.* [21] has been proven to be powerful approach in solving practical problems. This includes work done in pattern recognition, regression function estimation, nonlinear classification, time series prediction, density estimation and many more [12, 17]. It is a reformulation to standard Support Vector Machines (SVM) which lead to solve Karush-Kuhn-Tucker (KKT) systems. It is closely related to regularization networks and Gaussian process but additionally emphasizes and exploits primal-dual interpretations [7]. By adopting Structural Risk Minimization (SRM), its aim is to minimize an upper bound of generalization error which finally resulted in better generalization performance compared to other technique [6].

LSSVM is reported competent as it performs training and optimization relatively fast [11]. In solving a set of linear equations, LSSVM must first be trained to show its capability in producing a good prediction results. Nevertheless, during the training process, a problem encountered is the generalization ability of LSSVM, whether in classification or regression problem, is influenced by its hyper-parameters, which are regularization parameter, γ and kernel parameter, σ^2 . If hyper-parameters of the model are not well chosen, results produced will be not satisfied enough. Unfortunately, to overcome the problem, it requires effort in order to obtain the optimize hyper-parameters of LSSVM. In that matter, to get the ideal value of both

hyper-parameters, it clearly not an easy task since it has huge numbers of combinations. Hence, conventional factor selection which is man-made choice approach may not be a suitable technique since this technique depends on the expertise and experience of the researchers. Besides, it also involves tedious tasks by repeating countless of experiments. This situation may leads to some inconvenience issue such as selection randomness value in tuning the parameters. In addition, this approach is obviously time consuming and unsystematic [4, 5].

Due to that matter, most of the works rely on Cross-Validation (CV) for their optimization of LSSVM hyper-parameters (leave-one-out CV). However, CV is reported to be complicated in calculation and time consuming [8]. In addition, in terms of prediction accuracy, the results yielded by CV is in average [24]. To find the solution of this matter, many academic researchers have presented various approaches, particularly by applying Artificial Intelligence (AI) technique. Genetic Algorithm (GA), which is one of the most popular Evolutionary Algorithms (EAs) have been presented in [8, 17, 24] to solve problems in electrical, shipping and financial fields respectively. In a related work based on AI technique, Particle Swarm Optimization (PSO) has been proposed by [2] to predict the motion of ship and while artificial fish swarm algorithm has been studied by [4] to solve time series forecasting in electrical field. Besides the above approach, Artificial Bee Colony (ABC) was proposed by [18] in predicting gold price.

As a new nature inspired algorithm, ABC that has been proposed by Karaboga [13] has been proved of having strong ability to find global optimistic result. It is a new meta-heuristic population-based optimization technique motivated by the intelligent foraging behavior of honey bee swarms. ABC posses the advantages on memory, local search and the ability of self-enhancement in finding solution [19] where it leads to yield outstanding results in solving optimization issues. Thus, in order to achieve automatic selection of LSSVM hyper-parameters as well as improving the prediction accuracy, in this paper, a hybrid algorithm based on an enhanced ABC namely *e*ABC-LSSVM is applied in financial time series prediction. Simulation results verify the feasibility and validation the *e*ABC-LSSVM to financial time series prediction.

The remainder of this paper is organized as follows: Section 2 provides a brief introduction to LSSVM and ABC algorithm with its enhancement is describes in the following section. Section 4 discusses methodology implemented to design the prediction model. The results of proposed method with the comparison with standard ABC and CV method reports in section 5 and finally, conclusion is stated in section 6.

2. An Overview of Least Squares Support Vector Machines

LSSVM which stands for Least Squares Support Vector Machines reformulates the original SVM algorithm. It has been proposed by Suykens *et al.* [21] for the purpose to solve short term load prediction problems. LSSVM is reported to consume less computational effort in the huge-scale problem compared to standard SVM's. As a modified version of a standard SVM, LSSVM applies equality constraint instead of inequality constraint that has been used in SVM to obtain a linear set of equations [5]. Hence, it simplifies the complex calculation and easy to train [23]. In addition, in several real-cases demonstration, LSSVM has been reported to produce outstanding generalization performance with low cost in computational [24]. According to [3], the reformulation of standard SVM, LSSVM predicting model presented remarkable performance in simulation and practical results, compared to Radial Basis Function Neural Network (RBFNN) predictor and Back Propagation Neural Network (BPNN) predictor.

The standard framework for LSSVM is based on the primal-dual formulation. Given the dataset $\{x^i, y^i\}_{i=1}^N$, the aim is to estimate a model of the form [1]:

$$y(x) = w^T \phi(x) + b + e_i \quad (1)$$

Where $x \in R^n$, $y \in R$, and $\phi(\cdot): R^n \rightarrow R^{n_h}$ is a mapping to a high dimensional feature space. The following optimization problem is formulated [21]:

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

Subject to $y_i = w^T \phi(x_i) + b + e_i$, $i = 1, 2, \dots, N$.

With the application of Mercer's theorem [22] for the kernel matrix Ω as $\Omega_{ij} = K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ $i, j = 1, \dots, N$ it is not required to compute explicitly the nonlinear mapping $\phi(\cdot)$ as this is done implicitly through the use of positive definite kernel functions K . From the Lagrangian function:

$$\zeta(w, b, e; \alpha) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{i=1}^N e_i^2 - \sum_{i=1}^N \alpha_i (w^T \phi(x_i) + b + e_i - y_i) \quad (3)$$

Where $\alpha_i \in R$ are Lagrange multipliers. Differentiating equation 3 with w , b , e_i and α_i , the conditions for optimality can be described as follow:

$$\begin{cases} \frac{\partial \zeta}{\partial w} = 0 \rightarrow w = \sum_{i=1}^N \alpha_i \phi(x_i) \\ \frac{\partial \zeta}{\partial b} = 0 \rightarrow \sum_{i=1}^N \alpha_i = 0 \\ \frac{\partial \zeta}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i, i = 1, \dots, N \\ \frac{\partial \zeta}{\partial \alpha_i} = 0 \rightarrow y_i = w^T \phi(x_i) + b + e_i \end{cases} \quad (4)$$

By elimination of w and e_i , the following linear system is obtained:

$$\begin{bmatrix} 0 & I^T \\ y & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (5)$$

With $y = [y_1, \dots, y_N]^T$, $\alpha = [\alpha_1, \dots, \alpha_N]^T$. The resulting LS-SVM model in dual space becomes:

$$y(x) = \sum_{i=1}^N \alpha_i K(x, x_i) + b \quad (6)$$

In practice, the training of the LSSVM model involves an optimal selection of regularization parameter γ and kernel parameter σ^2 . Several kernel functions, viz. Gaussian Radial Basis Function (RBF) kernel, linear kernel and quadratic kernel is available. In this study, the RBF kernel is used which is expressed as:

$$K(x, x_i) = e^{-\frac{\|x-x_i\|^2}{2\sigma^2}} \quad (7)$$

Where σ^2 is a tuning parameter which associated with RBF function.

3. Artificial Bee Colony

For the purpose of obtaining optimized hyper-parameters of LSSVM, the recently developed Swarm Intelligent technique, ABC is integrated into LSSVM. ABC algorithm was proposed by Karaboga [13] for real parameter optimization. It is inspired by the intelligent behavior of honey bees. In the proposed model, the colony of artificial bees incorporated of three groups of bees: employed bees which associated

with specific food sources, onlooker bees which responsible to watch the dance of employed bees within the hive in order to choose a food source and scout bees which randomly search for food source. For onlookers and scouts, they are also known as unemployed bees [14]. Half of the colony is composed of employed bees and the rest are of the onlooker bees. The number of food sources/nectar sources is equal with the employed bees, which means that one employed bee is responsible for a single nectar source. The aim of the whole colony is to maximize the amount of nectar.

The duty of employed bees is to search for food sources (solutions). Later, the amount of nectars (solutions' qualities/fitness value) is calculated. Then, the information obtained is shared with the onlooker bees which are waiting in the hive (dance area). The onlooker bees decide to exploit a nectar source depending on the information shared by the employed bees. The onlooker bees also determine the source to be abandoned and allocate its employed bee as scout bees. For the scout bees, their task is to find the new valuable food sources. They search the space near the hive randomly.

In ABC algorithm, suppose the solution space of the problem is D -dimensional, where D is the number of parameters to be optimized. In this work, the parameters involved are γ and σ^2 . The fitness value of the randomly chosen site is formulated as follows:

$$fit_i = \frac{1}{(1 + obj.Fun_i)} \quad (8)$$

The size of employed bees and onlooker bees are both SN , which is equal to the number of food sources. For each food source's position, one employed bee is assigned to it. For each employed bee whose total numbers are equal to the number of the food sources, a new source is obtained according to equation 9:

$$v_{ij} = x_{ij} + \varphi_{ij}(x_{ij} - x_{kj}) \quad (9)$$

Where:

$i = 1, 2, \dots, SN$.

$j = 1, 2, \dots, D$.

φ = a random generalized real number within the range $[-1, 1]$.

k = is a randomly selected index number in the colony.

After producing the new solution, $v_i = \{x'_{i1}, x'_{i2}, \dots, x'_{iD}\}$, it is compared to the original solution $v_i = \{x_{i1}, x_{i2}, \dots, x_{iD}\}$. If the new solution is better than previous one, the bee memorizes the new solution; otherwise she memorizes the former solution. The onlooker bee selects a food source to exploit with the probability:

$$p_i = \frac{fit_i}{\sum_{j=1}^{SN} fit_j} \quad (10)$$

Where fit_i is the fitness of the solution v . SN is the number of food sources positions. Later, the onlooker bee searches a new solution in the selected food source site by equation 9, the same way as exploited by employed bees. In SB phase, if the fitness of a found food source hasn't been improved for a given number of trial (denoted by limit), it is abandoned. This action represents the negative feedback in ABC algorithm and the EB of that food source becomes a SB and makes a random search by equation 11:

$$x_{id} = x_d^{\min} + r(x_d^{\max} - x_d^{\min}) \quad (11)$$

Where:

r = a random real number within the range $[0,1]$.

x_d^{\min} and x_d^{\max} = the lower and upper borders in the d^{th} dimension of the problems space.

Basic steps of ABC algorithm are as follows:

1. Initialize the food source positions (population).
2. Each employed bees is assigned on their food sources.
3. Each onlooker bee select a source base on the quality of her solution, produces a new food source in selected food source site and exploits the better source.
4. Decide the source to be cast aside and assign its employed bee as scout for discovering new food sources.
5. Memorize the best food source found so far.
6. If requirement are met, output the best solution, otherwise repeat steps 2-5 until the stopping criterion is met.

• Enhanced Artificial Bee Colony

As a new member of meta heuristic family, ABC has proven to be an effective approach in solving optimization problem. As compared to GA and Ant Colony Optimization (ACO), ABC revealed its capability in producing remarkable performance [20]. However, sometimes, ABC operates too well. If precaution step is not considered, ABC model inclined to converge too fast and this may lead to local minima. This might happen if the area explored by the model is not a desired area, namely global minima area. Thus, an improvement in ABC is made by adopting mutation approach [10]. By including mutation strategy in ABC, the model is induced to explore other areas in order to look for global minima rather than local minima.

In ABC-LSSVM [18], if the generated parameter value is out of boundaries, it is automatically moved onto the boundaries. Nevertheless, in *e*ABC-LSSVM, some modification is made. When the unwanted situation aforementioned occurred, instead of forcing the parameter value to the boundary, mutation strategy is introduced. This operation is executed by

multiplying the generated random number with the range of boundary that has been determined. In this study, the boundaries are set to the range of between [1, 1000]:

$$new_param = (ub - lb) * rand_num \quad (12)$$

Where:

new_param = new parameter
 $rand_num$ = random number
 ub = upper bound
 lb = lower bound

Figure 1 shows the simplified form of the proposed prediction model while the flow of eABC algorithm with the modification that has been made is illustrated in Figure 2. From Figure 2, it can be seen that the mutation approach is applied in both employed and onlooker bees phases prior to producing new food solution.

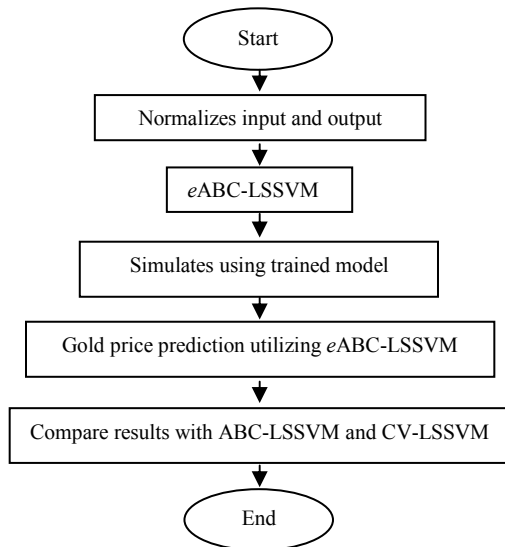


Figure 1. Flow of LSSVM.

4. Methodology

In this section, the research data utilized, pre-process technique applied and experimental setup are described.

4.1. Data Preparation and Pre-Process

Data selected for a prediction model should represent variables that affect the prediction model. In this paper, to demonstrate the efficiency of proposed model, a financial time series was utilized, namely gold price time series. For the prediction purposes, besides the world gold prices time series itself, other commodity prices which are daily spot prices of West Texas Intermediate (WTI) Cushing, Oklahoma, such as crude oil and commodity-relevant macro financial variables, namely Euro, Yen and GBP to USD exchange rate data sets were also employed as independent variables. The feature of WTI is included as an input due to its significant influence to the fluctuation of many commodities markets, including gold [26]. Euro Area

and Japan currencies are chosen since they are part of the leading trading partners of the US [9]. All time series were provided by The London Bullion Market Association (LBMA) website, US Energy Information Administration (EIA) and OANDA website.

The time series frequency is daily, from January 2008 until July 2011. The time series was divided into three parts, namely training, validation and testing set, where 70% of the data were three used as the training set, which is from day 1 to day 905. Another 15% is for validating and the balance of 15% for testing set.

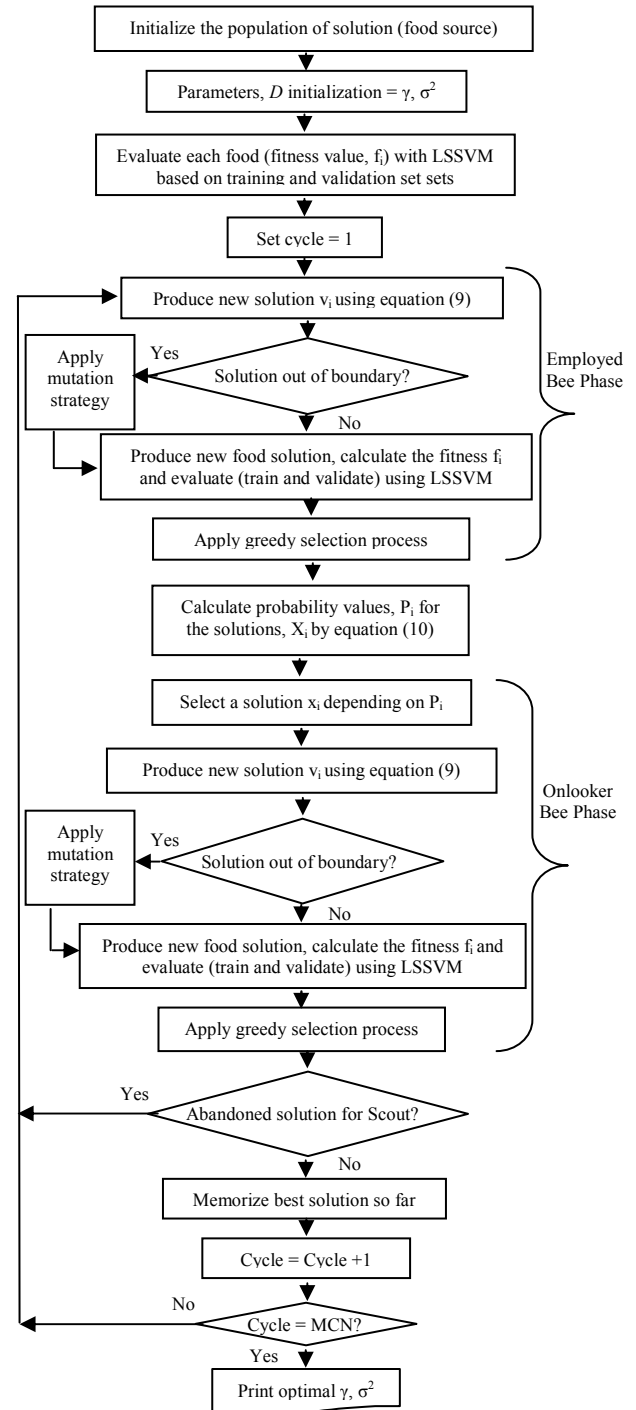


Figure 2. Flow of eABC-LSSVM.

Both the testing and validating data sets are completely different from the training sets and have

been isolated from the training process. Table 1 shows the number of samples used in training, validation and testing processes while Table 2 indicates the variables assigned to features involved.

Table 1. Number of samples for training, validation, and testing.

Data Types	Samples (Days)
Training	1-905
Validation	906-1099
Testing	1100-1293

Table 2. Assigning input and output variables.

	Input	Variable	Output
Gold Price Prediction	Gold price	gold	gold
	Crude oil price	wti	
	Major currencies (Euro, Yen and GBP)	eur yen gbp	

4.2. Normalization Process

The successful of machine learning algorithms rely on the quality of the data they operate on [16]. Thus, data transformation such as normalization may improve the prediction accuracy and efficiency of mining algorithm [1].

In this study, data collected is from multiple sources. Data mining seeks to discover unrecognized associations between data items in related databases. Therefore, data normalization technique is applied to all input and output vectors before sending to the prediction model for learning. For time series prediction, external normalization is most widely applied [25], where all the involved data are normalized into specific range. This is an important stage to make sure that the data are not overwhelmed by each other in terms of distance measure. In this paper, min max normalization was applied. The formula used for the min max normalization is as equation 13:

$$v' = \frac{(v - \min_a) / (\max_a - \min_a)}{(\text{new-max}_a - \text{new-min}_a) + \text{new-min}_a} \tag{13}$$

In min max normalization, it maps a value of v of A to v' in the range [new-max_a, new-min_a] by solving the equation above. Data tabulated in Tables 3 and 4 show the samples of original data used in this work and the normalized data.

Table 3. Samples of original input for gold price prediction.

Gold	Wti	Gbp/Usd	Eur/Usd	Yen/Usd
858.375	95.08	1973.6	1473.8	9.2
873.375	96.43	1971.4	1471.3	9.2
882.425	95.64	1973.1	1470.5	9.1

Note: Currencies are based on GBP/EUR/YEN1000

Table 4. Samples of input for gold price prediction after min max normalization.

Gold	Wti	Gbp/Usd	Eur/Usd	Yen/Usd
0.1734	0.5633	0.9130	0.6974	0.0789
0.1901	0.5751	0.9097	0.6912	0.0789
0.2001	0.5682	0.9123	0.6892	0.0526

4.3. Performance Evaluation Metric

In order to justify the results, it is important to choose the appropriate evaluation indicators. In this work, for prediction evaluation measurement, a quantitative evaluation is used, namely Mean Absolute Percentage Error (MAPE). It is one of the most common evaluations metric applied in prediction [15]. In MAPE, the errors are measured relative to the data values. The formula of MAPE is shown in equation 14:

$$MAPE = \frac{1}{n} \left[\sum_{i,j=1}^n \left| \frac{Y_i - P_i}{Y_i} \right| \right] \tag{14}$$

Where:

$i, j = 1, 2, \dots, n.$

Y = Actual value.

P = Prediction value.

Another evaluation metric utilized is prediction accuracy, which is shown in equation 15:

$$PA = \text{mean}(100\% - (MAPE \times 100)) \tag{15}$$

Where PA = Prediction Accuracy.

5. Experimental Results

In this section, the properties of control parameters for ABC is as tabulated in Table 5 while empirical finding from experiments conducted is presented in Table 6. The final optimal value for γ and σ^2 in eABC-LSSVM is set to 912.3238 and 13.7854 respectively. Both values have produced good results for eABC-LSSVM in the prediction model. The prediction result generated by eABC-LSSVM offers better MAPE as compared to the other two competitors which is 0.6885, which reflects the prediction accuracy with 99.3115%. This is followed by MAPE yielded by ABC-LSSVM, which is 1.0179 with $\gamma=691.7331$ and $\sigma^2=8.4373$ and finally MAPE obtained by CV-LSSVM, which is 3.8187, with both γ and σ^2 is set to 59.245 and 2.41296, respectively.

Table 5. Control parameters for ABC.

Parameters	Value
Number of colony size (Employed + Onlooker)	20
Food number	10
Limit	2

Table 6. eABC-LSSVM vs. ABC-LSSVM vs. CV-LSSVM in gold price prediction.

	eABC-LSSVM	ABC-LSSVM	CV-LSSVM
γ	912.3238	691.7331	59.245
σ^2	13.7854	8.4373	2.41296
MAPE (%)	0.6885	1.0179	3.8187
Prediction Accuracy (%)	99.3115	98.9821	96.1813

Figure 3 depicts the comparison between target value and predicted value by eABC-LSSVM, ABC-

LSSVM and CV-LSSVM in gold price prediction while the comparison in term of convergence between e ABC-LSSVM and ABC-LSSVM is illustrated in Figure 4. The efficiency of proposed model can be seen from the narrow span showed in the graph, where the predicted value by e ABC-LSSVM is more accurate than ABC-LSSVM and CV-LSSVM. In addition, the proposed model also offers better convergence performance as compared to standard ABC-LSSVM.

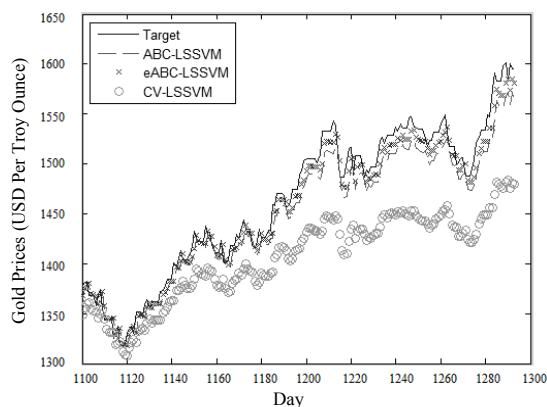


Figure 3. Target vs. ABC-LSSVM vs. e ABC-LSSVM vs. CV-LSSVM for gold price prediction.

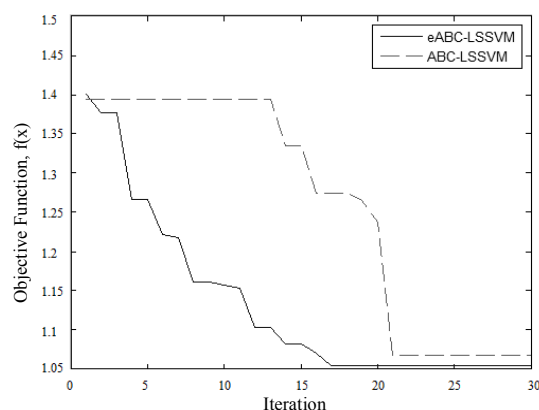


Figure 4. Convergence performance of ABC-LSSVM vs. e ABC- for gold price prediction.

6. Conclusions

In this study, the combination of machine learning technique with enhanced optimization approach is presented through e ABC-LSSVM to predict gold price. Data normalization was first performed to improve the chance in increasing accuracy and efficiency of data mining algorithm. By utilizing mutation approach, the capability of existing ABC approach is enhanced to better optimize LSSVM's hyper parameters. Final output clearly shows that the use of e ABC is of a better solution than the ABC and CV in tuning LSSVM parameters, particularly in predicting the gold prices. By increasing the capability of standard ABC in escaping local minima, the e ABC has better global search capability in achieving global minima, hence producing satisfied prediction accuracy. As a

consequence, the proposed e ABC-LSSVM can further be explored in other domain of study.

References

- [1] Al-Shalabi L., Shaaban Z., and Kasasbeh B., "Data Mining: A Preprocessing Engine," *Journal of Computer Science*, vol. 2, no. 9, pp. 735-739, 2006.
- [2] Bo Z. and Aiguo S., "LSSVM and Hybrid Particle Swarm Optimization for Ship Motion Prediction," in *Proceedings of the International Conference on Intelligent Control and Information Processing*, Dalian, China, pp. 183-186, 2010.
- [3] Chen Q., Wu Y., and Chen X., "Research on Customers Demand Forecasting for E-business Web Site Based on LS-SVM," in *Proceedings of the International Symposium on Electronic Commerce and Security*, Guangzhou City, China, pp. 66-70, 2008.
- [4] Chen X., Wang J., Sun D., and Liang J., "Time Series Forecasting Based on Novel Support Vector Machine Using Artificial Fish Swarm Algorithm," in *Proceedings of the 4th International Conference on Natural Computation*, Jinan, China, pp. 206-211, 2008.
- [5] Cheng G., Guo R., and Wu W., "Petroleum Lithology Discrimination Based on PSO-LSSVM Classification Model," in *Proceedings of the 2nd International Conference on Computer Modeling and Simulation*, Sanya, China, pp. 365-368, 2010.
- [6] Djemili R., Bourouba H., and Korba A., "A Combination Approach of Gaussian Mixture Models and Support Vector Machines for Speaker Identification," *the International Arab Journal of Information Technology*, vol. 6, no. 5, pp. 490-497, 2009.
- [7] Espinoza M., Suykens J., and Moor B., "Fixed-Size Least Squares Support Vector Machines: A Large Scale Application in Electrical Load Forecasting," *Computational Management Science*, vol. 3, no. 2, pp. 113-129, 2006.
- [8] Fu H., Liu S., and Sun F., "Ship Motion Prediction Based on AGA-LSSVM," in *Proceedings of the International Conference on Mechatronics and Automation*, Xi'an, China, pp. 202-206, 2010.
- [9] Hammoudeh S., Sari R., and Ewing B., "Relationships Among Commodities and with Financial Variables: A New Look," *Contemporary Economic Policy*, vol. 27, no. 2, pp. 251-264, 2008.
- [10] Haupt R. and Haupt S., *Practical Genetic Algorithms*, Wiley-Interscience Publication, New York, 1998.

- [11] Hou L., Yang Q., and An J., "An Improved LSSVM Regression Algorithm," in *Proceedings of the International Conference on Computational Intelligence and Natural Computing*, Wuhan, China, vol. 2, pp. 138-140, 2009.
- [12] Kamruzzaman J., Sarker R., and Ahmad I., "SVM Based Models for Predicting Foreign Currency Exchange Rates," in *Proceedings of the 3rd IEEE International Conference on Data Mining*, pp. 557-560, 2003.
- [13] Karaboga D., "An Idea Based on Honey Bee Swarm for Numerical Optimization," *Technical Report*, Erciyes University, 2005.
- [14] Karaboga D., Gorkemli B., Ozturk C., and Karaboga N., "A Comprehensive Survey: Artificial Bee Colony (ABC) Algorithm and Applications," *Artificial Intelligence Review*, vol. 41, no. 138, 2012.
- [15] Klimberg R., Sillup G., Boyle K., and Tavva V., "Forecasting Performance Measures-What are Their Practical Meaning?," *Advances in Business and Management Forecasting*, vol. 7, pp. 137-147, 2010.
- [16] Kotsiantis S., Kanellopoulos D., and Pintelas P., "Data Preprocessing for Supervised Learning," *the International Journal of Computer Science*, vol. 1, no. 2, pp. 1306-4428, 2006.
- [17] Mustafa M., Khalid S., Aliman O., Sulaiman M., Rahim S., and Shareef H., "Tracing the Real Power Transfer of Individual Generators to Loads using Least Squares Support Vector Machine with Continuous Genetic Algorithm," in *Proceedings of the International Conference on Electrical, Control and Computer Engineering*, Pahang, Malaysia, pp. 76-81, 2011.
- [18] Mustafa Z. and Yusof Y., "Optimizing LSSVM Using ABC for Non-Volatile Financial Prediction," *Australian Journal of Basic and Applied Sciences*, vol. 5, no. 11, pp. 549-556, 2011.
- [19] Shi X., Li Y., Li H., Guan R., Wang L., and Liang Y., "An Integrated Algorithm Based on Artificial Bee Colony and Particle Swarm Optimization," in *Proceedings of the 6th International Conference on Natural Computation*, Shandong, China, vol. 5, pp. 2586-2590, 2010.
- [20] Shokouhifar M. and Abkenar G., "An Artificial Bee Colony Optimization for MRI Fuzzy Segmentation of Brain Tissue," in *Proceedings of the International Conference on Management and Artificial Intelligence*, Indonesia, vol. 6, pp. 6-10, 2011.
- [21] Suykens J., Van-Gestel T., De-Brabanter J., De-Moor B., and Vandewalle J., *Least Squares Support Vector Machines*, World Scientific Publishing, Singapore, 2002.
- [22] Vapnik V., *The Nature of Statistical Learning Theory*, Springer, New York, 1995.
- [23] Yang L., "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*, Curitiba, Brazil, pp. 1-5, 2009.
- [24] Yu L., Chen H., Wang S., and Lai K., "Evolving Least Squares Support Vector Machines for Stock Market Trend Mining," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 1, pp. 87-102, 2009.
- [25] Zhang G., Patuwo B., and Hu M., "Forecasting with Artificial Neural Network: The State of Art," *the International Journal of Forecasting*, vol. 14, no. 1998, pp. 35-62, 1998.
- [26] Zhang Y. and Wei Y., "The Crude Oil Market and the Gold Market: Evidence for Cointegration, Causality and Price Discovery," *Resources Policy*, vol. 35, no. 3, pp. 168-177, 2010.



Zuriani Mustaffa received her Bachelor of Science (Computer) from the Universiti Teknologi Malaysia in 2004 and Master of Science (Information Technology) from the Universiti Utara Malaysia in 2011. She is currently pursuing her PhD degree in information technology, from Universiti Utara Malaysia. Her research interest is in computational intelligence, specifically in swarm intelligence and machine learning.



Yuhanis Yusof is a senior lecturer in the School of Computing, the Universiti Utara Malaysia, Malaysia. She obtained her PhD in computer science from Cardiff University, UK. She also holds MSc degree in computer science and Bachelor in information technology. Her research interest is broadly in data analysis and management for large scale computing. This includes data mining (discovering patterns of interest from data), data warehousing, information retrieval, digital libraries and software reuse. She is also involved in the area of computational intelligence, particularly in machine learning and swarm computing.