A REVIEW ON OPTIMIZATION OF LEAST SQUARES SUPPORT VECTOR MACHINE FOR TIME SERIES FORECASTING

Yuhanis Yusof[1] and Zuriani Mustaffa[2]

[1]School of Computing, Universiti Utara Malaysia, Malaysia
[2]Faculty of Computer Systems & Software Engineering, Universiti Malaysia Pahang, Malaysia

ABSTRACT

Support Vector Machine has appeared as an active study in machine learning community and extensively used in various fields including in prediction, pattern recognition and many more. However, the Least Squares Support Vector Machine which is a variant of Support Vector Machine offers better solution strategy. In order to utilize the LSSVM capability in data mining task such as prediction, there is a need to optimize its hyper parameters. This paper presents a review on techniques used to optimize the parameters based on two main classes; Evolutionary Computation and Cross Validation.

KEYWORDS

Least Squares Support Vector Machine, Evolutionary Computation, Cross Validation, Swarm Intelligence

1. INTRODUCTION

The Least Squares Support Vector Machines (LSSVM) [1] which is an extension of Support Vector Machines (SVM) [2] has contributed significant impact in machine learning community. LSSVM has been recognized as one of the tool in solving various data mining tasks which include prediction, classification and many others [3].Nonetheless, besides its diversity in application, the capability of LSSVM is highly dependent on the value of its hyper-parameters, namely regularization parameter, γ and kernel parameter, σ^2[4].

Time series prediction is of interest in various fields. It is important due to prediction of future values could be a vital input for current planning and decision making [5]. To realize the time series prediction, historical data is employed to generate future values, i.e. to make prediction[6]. Due to the importance of time series prediction, numerous approaches have been presented, and this includes in price prediction[7,8]. In literature, the price prediction modelswere presented in various horizon or time scales; short term, mediumor long term[9]. Thus, different data frequency has been employed depending on the suitability viz. daily, weekly, monthly or annually[10,11]. This paper reviews existing work that optimizes the aforementioned LSSVM parameters and the discussion is divided into two sections; Optimization based on Cross Validation and Optimization based on Evolutionary Computation.

2. OPTIMIZATION OF LSSVM USING CROSS VALIDATION

Researchers in [12] highlight the importance of primary energy consumption prediction by presenting a prediction model utilizing LSSVM. Several factors that were chosen to be fed to the prediction models include industrial index and population. Empirical data sets were from 1990 to 2008, which was limited for the case in China. For evaluation purposes, two performance criteria were used namely Absolute Percentage Error (APE) and Mean Absolute Percentage Error (MAPE). For parameter tuning, Cross Validation (CV) technique was employed. By using CV, each training set, TR, is divided into X partitions in equal size (TR_1, TR_2, …,TR_X). For each run, the LSSVM is trained X times and one of X is set as testing set. Later, the trained model is used in the testing set. Finally, by using the obtained parameters value, the average accuracy of X trials is calculate to estimate the final accuracy. The CV-LSSVM technique achieved the predetermined condition where the accepted error range is within 5%. In addition, as the researchers employed yearly samples, it shows that LSSVM is capable to deal with small samples data sets as compared to Artificial Neural Network (ANN) which frequently needs more data for training purposes [13].

In 2008, Liu and Wang presented a prediction technique that is based on LSSVM to predict foreign exchange rate market [14]. In order to determine the optimized hyper-parameters of LSSVM, CV approach was applied. For experimental purposes, four pairs of exchange rates time series data were employed. The samples comprise of time series data from 2003-2007, which consist of 1196 observations. By using 86:14 data proportion for training and testing respectively, the efficiency of proposed technique is evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and MAPE. However, even though the CV-LSSVM is capable to provide promising results, the absence of comparison with other comparable prediction techniques made the experiment incomplete.

In related work, an Empirical Mode Decomposition (EMD) was integrated with LSSVM in dealing with exchange rate prediction [15].The EMD was utilized to decompose the foreign exchange rates into several International Monetary Fund (IMF) components and one residual component. The LSSVM then performs the prediction of the IMFs and residual value independently. Comparing the result against those produced by EMD-ARIMA, single LSSVM and single ARIMA, the EMD-LSSVM outperforms other experimented prediction techniques based on several criteria, which includes MAE, MAPE, Correct Uptrend (CP) and Correct Downtrend (CD). For parameter tuning, the CV approach was employed.

Progressing further, in [16], short term prediction of meteorological time series data has been presented using CV-LSSVM. Prior to training task, three pre-processing techniques were performed namely data filtering which is to remove redundant components from the data sets, removing outliers from the data sets and finally data normalization. The evaluation of the presented technique is based on several statistical metrics which includes MAPE and RMSE. Comparison againts various ANN based predictor showed that the performance of CV-LSSVM is at par with Probabilistic Neural Network (PNN).

Meanwhile, [17] employed LSSVM to predict short-term power load in Inner Mongolia. The case study used one year load series from that region, and later divided them into 90:10 for training and testing respectively. For parameter tuning, CV technique was utilized. The results derived from this approach is regarded as efficient for the problem under study. However, no comparison with other technique was conducted.

A study in [18] integrated the Empirical Mode Decomposition (EMD) with LSSVM in predicting future wind speed and output power. The purpose of applying EMD is similar as described in [15] which is to decompose the data sets utilized into a series of different scales of IMF. By employing daily average wind speed from March to May 2006, the result obtained by the proposed model is compared against several techniques which include single LSSVM. The hyper-parameters of LSSVM are tuned based on CV approach. The comparison was made by taking into account the value produced in terms of MAPE and RMSE by each identified prediction model. Final results indicated that the EMD-LSSVM outperformed the other experimented techniques by producing lower MAPE and RMSE.

In related work [19], demonstrated a CV-LSSVM for air temperature prediction. This study utilized monthly average temperature from 1951 to 2003. From that sample, only year 2003 data is allocated for testing set. The findings of the experiment conducted were evaluated using RMSE, Relative Error (RE) and prediction accuracy and involved another two prediction techniques, namely single LSSVM and Radial Basis Function Neural Network (RBFNN). Final results indicated that RMSE produced by EMD-LSSVM was 0.7202, which was the smallest among the experimented prediction techniques.

Meanwhile, in [20], the CV-LSSVM was presented in prediction of flashover voltage of insulators. In the study, several comparisons were performed between different types of kernels, which include RBF kernel and Polynomial kernel. The results indicated that Radial Basis Function (RBF) kernel yield lower RMSE. Meanwhile, the comparison against ANN is in favor to CV-LSSVM.

Previously, the application of LSSVM in a Bayesian evidence framework was examined by [21]. In the presented study, two different cases were investigated, namely prediction for U.S. short term interest rate and daily index closing price return. Prior to training, normalization technique was performed to all data sets. Upon completing the experiments, the development of proposed model is proven to be successful as an improvement is made in MSE, MAE and Negative Log Likehood (NLL) as compared to traditional methods, which includes autoregressive model.In [22], a hybridization of Group Method Data Handling (GMDH) with LSSVM was presented to predict electrical energy consumption. GMDH is the ANN based model which finds the optimal model based on inductive sorting out approach. For LSSVM parameter selection, CV technique was utilized. When evaluated based on MAPE, the presented technique offers lower MAPE as compared to the comparison algorithms, which includes single GMDH and LSSVM.

As discussed above, even the utilized CV has provided promising results for the context of study, however, without comparing CV with other techniques in the study conducted [14,23,24], proving the efficiency of the presented method is therefore difficult. In addition, CV is revealed to be time consuming [25] and tends to produce high error rates[26]. Due to that matter, an application of Evolutionary Computation (EC) algorithm, which includes Evolutionary Algorithm (EA) and Swarm Intelligence (SI) techniques, become research interest among academic community in optimizing LSSVM hyper-parameters which is presented in the next section. A summary of this section is tabulated in Table 1.

Table 1: Time Series Prediction Technique using LSSVM Optimized by CV

Author	Research Area	Independent Variable	Method	Data Frequency	Evaluation Metric
Wang & Li (2010)	Primary energy consumption prediction	Gross Domestic Product (GDP), population and others	CV-LSSVM	Annually	APE and MAPE
Liu & Wang (2008)	Exchange rates prediction	JPY/USD, CAD/USD, USD/GBP and USD/EUR	CV-LSSVM	Daily	MAE, MAPE AND RMSE
Lin, et al.(2012)	Exchange rates prediction	USD/NTD, JPY/NTD, and RMB/NTD	CV-EMD-LSSVM	Daily	MAE, MAPE, CP, CD
Mellit et al.(2013)	Short-term meteorological time series data prediction	Solar irradiance, air temperature, relative humidity and pressure	CV-LSSVM	Daily	R^2, RMSE, Mean Bias Error (MBE), MAPE
Wu &Niu (2009)	Short-term load prediction	Load series in Inner Mongolia	CV-LSSVM	Daily	MSE
Wang & Li (2009)	Wind speed and output power prediction	Average wind speed	EMD-LSSVM	Daily	MAPE and RMSE
Wang, et al.(2010)	Air temperature prediction	Average temperature	EMD-LSSVM	Monthly	RMSE, RE
Gencoglu&Uyar (2009)	Flashover voltage of insulators prediction	Insulator height, insulator diameter, and several others	-	CV-LSSVM	RMSE, R^2, Mean Error Function (MFE)
Van Gestel, et al.(2001)	Financial time series prediction	Interest rate, closing index prices	Weekly	CV-LSSVM	MSE, NLL and MAE
Ahmad, et al. (2012)	Electrical energy consumption prediction	Electrical energy consumption time series data	-	CV-GMDH-LSSVM	MAPE

2.1 Optimization of LSSVM using Evolutionary Computation

The Evolutionary Computation (EC) comprises of Evolutionary Algorithm (EA) and SI techniques[27]. The EA is inspired by the natural evolution phenomenon while the imitation of insects or animals is featured in SI technique[28]. Figure 1 illustrates the EC classes.

Evolutionary Computation (EC)

Evolutionary Algorithm (EA)

Swarm Intelligence (SI)

Genetic Algorithm (GA),
Evolutionary Programming (EP)
Differential Evolution (DE)

Artificial Bee Colony (ABC)
Fish School (FS)
Particle Swarm Optimization (PSO)
Artificial Fish Swarm Algorithm (AFSA)
Ant Colony Optimization (ACO)
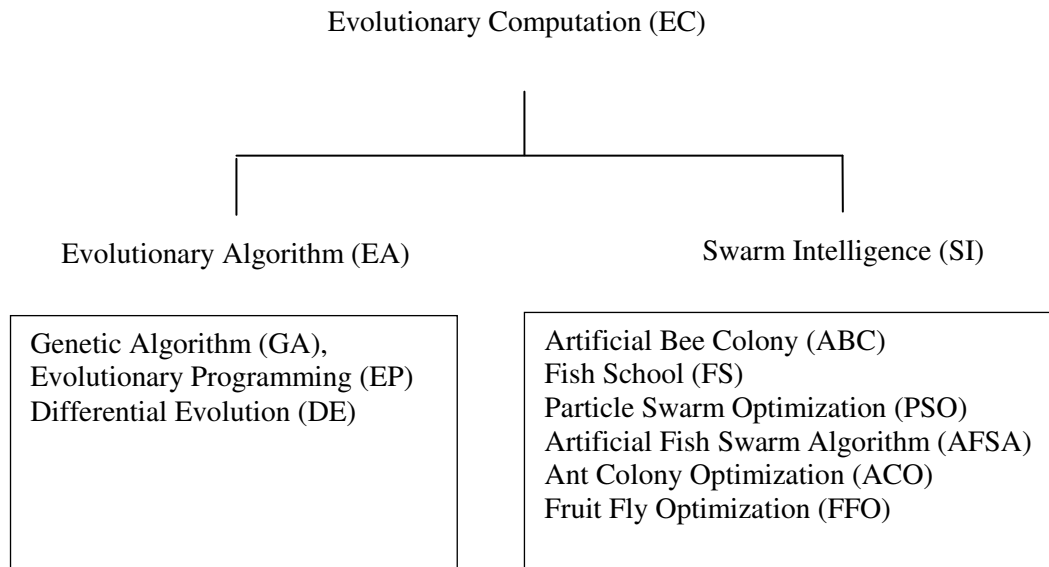Fruit Fly Optimization (FFO)

Figure 1: Classes of Evolutionary Computation

Categorized under EA group, Genetic Algorithm (GA) is the most prominent and widely used technique as compared with other techniques within the same domain[28,29]. It was introduced by John Holland and later was popularized by David Goldberg[30]. It models genetic selection and imitate the natural evolution phenomenon to solve numerical optimization problems. The GA consists of five main components viz. random number generator, fitness evaluation, reproduction, crossover and mutation operations. The last three components are categorized as genetic operator. In GA, the possible solutions are represented by chromosomes.

The hybrid technique of GA-LSSVM has been demonstrated to predict transmission loss in deregulated power system [31]. In the study, the application of GA-LSSVM is used to learn which generator is responsible for the losses by determining the optimal value for LSSVM hyper-parameters. The data was broken down into three sets; 57% of samples are used for training, 14% are set for validation and the rest 29% is reserved for testing. Simulations were performed and the results obtained showed satisfactory result. Similar approach also has been tested by the authors to predict reactive power which is compared against result produced by ANN[32]. The results obtained suggested that both methods produced similar outputs; nonetheless GA-LSSVM possesses an advantage in terms of computational time.

Researchers in [26] presented the GA-LSSVM for stock market prediction. In this study, the first GA was utilized for feature selection while the second GA is for parameter tuning. The proposed model is compared against the results produced by several techniques which include conventional BPNN, ARIMA, SVM and variants LSSVM. From the experiment conducted, the GA-LSSVM is proven to be able to present some selection models that are easier to clarify by applying small number of features. However, for certain cases, the presented model was unable to outperform the other comparison algorithms.

In the maritime field, ship motion prediction by utilizing Adaptive GA (AGA) with LSSVM has been presented by [33]. In the study, by referring to the fitness values of individuals, the probability of crossover and mutation are adaptively adjusted. To prove the efficiency of the AGA-LSSVM, comparison with the BPNN and LSSVM were made. From the results obtained, it clearly shows that AGA-LSSVM excels by yielding lower MAE as compared to the experimented techniques. However, it would be appropriate if the comparison against GA-LSSVM is considered as to see the difference between the newly GA with the original one.

Meanwhile, in [34], they examined the capability of GA-LSSVM in power spot prices prediction. Prior to training, the data was pre-processed by applying Min Max Normalization. The performance of the model was measured based on MAE and MAPE. After completing experiment process, the average of MAE and MAPE produced by GA-LSSVM yielded lower value as compared to value produced by ANN. In [3], the comparison between the GA-LSSVM, standard SVM and LSSVM to predict dissolved gas concentration was performed. The findings of the study indicated that the GA-LSSVM provided better results than the other two approaches.

In [29], the GA-LSSVM was proposed for conductivity and tensile strength prediction for carbon fiber composites. In this study, the predictability of proposed model was guide based on two indices, namely Learning Mean Square Error (LMSE) and Generalization Mean Square Error (GMSE). To further assess the superiority of GA-LSSVM, comparisons against Principal Component Analysis-GABPNN (PCS-GABPNN) was made. Based on the results obtained, the GA-LSSVM is capable to produce better generalization performance without having to use any data pre-treatment as required in the identified competitor. In [35], the application of GA-LSSVM

was proposed to predict the condition of hydro turbine generating units (HGUs). Good prediction model for the context study is beneficial in providing backups to predictive maintenance system in mechanical equipment fault diagnosis. As compared to CV-LSSVM and BPNN, GA-LSSVM performed slightly better than CV-LSSVM while BPNN performed dismally due to over fitting problem.

In 2005, LSSVM optimized by GA has been used to deal with prediction of chaotic time series data[36]. Tested on benchmark time series data, both training and testing set were divided equally. The comparison against CV-LSSVM suggested that GA-LSSVM is capable to produced lower error rate, relative to the RMSE. Further development of GA-LSSVM has been presented by [37] (2013) in solving network flow prediction. In the study, the modification is put forward to GA, where Self-adaptive GA is introduced. The modification involved the standard operation of crossover and mutation. Using hourly network access flow, the data sets were first normalized using Min Max Normalization. The conducted experiment indicated that Self-adaptive GA-LSSVM is superior over the other identified approach. Nevertheless, the comparison with the standard GA-LSSVM was absent in the study.

Introduced by Kennedy and Eberhart in 1995 [38], PSO is considered as a well-known SI technique which was inspired by social behavior of birds flocking or fish schooling. The application of PSO can be seen broadly utilized in solving various optimization issues [28]. In this technique, it is initialized with a group of random particles which represent possible solutions, and then searches for global optimal by updating generations. As compared to GA, there is no evolution operator such as crossover and mutation in PSO, which make it simpler in implementation. However, the choice in parameter selection such as acceleration constant and weight still need to be considered[39].

The efficiency of PSO-LSSVM has been demonstrated by [4] in critical heat flux prediction (CHF). Applied in nuclear science field, accurate prediction of CHF is vital not only for safety but also in designing nuclear reactors. For that matter, several relevant inputs are considered for training and later, the samples of data sets are divided into training and testing. The performance of the proposed method is guided based on three indices, viz. the R^2, Mean Relative Error (MRE) and RMSE. Compared against ANFIS, the result obtained is in favor to the PSO-LSSVM.

Prediction for water quality utilizing hybrid of PSO with LSSVM was studied by [40]. In this study, 3256 data points of water figures from Lixiu River water database were employed as empirical data. The results obtained by the presented technique are compared against the results produced by using CV-LSSVM, BPNN, ARIMA and GML. The incapability of CV in this study supports the results obtained in [41] and confirms the fact stated in previous the studies[25]. However, while the PSO-LSSVM outperforms CV-LSSVM, BPNN and ARIMA, it loses to GML as prediction accuracy presented by the GML technique is higher.

Unlike the previous study of PSO-LSSVM [4, 40], in [42], the dynamic inertia weight PSO (WPSO) is employed to optimize LSSVM hyper-parameters. The optimal value of parameters obtained is then used by LSSVM for prediction task. Upon completing simulation task, the RE, which is selected as performance evaluation metric of the study indicated that the PSO-LSSVM outperformed both experimented prediction techniques, viz. BPNN and LSSVM. Since the PSO used in this study has been modified, a comparison with the standard PSO needs to be performed.

While the work in [3] demonstrated GA-LSSVM for gas prediction, [43] examined the capability of PSO-LSSVM for the similar task. In this study, dissolved gas data from several electric power companies in China were employed as empirical data. As to improve the generalization of LSSVM, the raw data were first normalized before training task. For comparison purposes, four single prediction models were chosen, including three ANN based prediction models and SVR. Upon completing the simulation task, the PSO-LSSVM was concluded to be superior than the identified techniques relative to MAPE and R^2. However, despite the good performance of the model, the author admitted that the produced results might be unreliable due to the insufficient size of actual data.

Experimented on chaotic time series data, [44] presented the use of PSO-LSSVM. For empirical purposes, two benchmark data sets were employed namely Mackey–Glass time series data and Lorenz time series data. As compared to fuzzy prediction based on SVM, the PSO-LSSVM showed superior performance for both cases.

As [33] presented AGA-LSSVM for ship motion prediction, [45] demonstrated the hybridization of PSO with LSSVM for the similar context of study. Taking the similar approach as in [33], the optimization technique, which is PSO, is modified as to enhance the generalization performance of LSSVM in prediction. The modification of PSO is done by integrating it with Simulated Annealing (SA), termed ad Hybrid Particle Swarm Optimization (HPSO). Tested on a real data, the performance of proposed method which was guided by RMSE produced an encouraging result. However, the evaluation does not involve any other prediction techniques.

Carbon price prediction using hybrid ARIMA and LSSVM has been demonstrated by [46]. For optimization of hyper-parameters of interest, PSO algorithm was employed. However, the PSO was only used to tune the LSSVM hyper-parameters, while the ARIMA parameters are derived empirically. The research data employed was daily carbon price which covers the period from 2005 to 2010. The performance of proposed model was compared against several single and hybrid techniques, which includes ARIMA-ANN and also LSSVM and the evaluation is based on RMSE. Findings of the study suggested that the presented model demonstrate better predictive capability by producing lower error.

Besides GA and PSO, another EC algorithm which includes Artificial Fish Swarm Algorithm (AFSA) [47,48], Ant Colony Optimization (ACO) [49], Differential Evolution (DE) [50] and Gravitational Search Algorithm (GSA)[51] were also presented in literature as one of the optimization tool for LSSVM.

AFSA which was first proposed by [48] and is inspired by the intelligent swarming behavior of fish. In AFSA, after initialization of population, the algorithm will go through the selection of the best behavior to be executed, which involved between swarm and follow behavior. The behavior which produced the minimum result will be chosen for the next stage. Even though AFSA possesses an advantage in avoiding local minimum, it suffers from imbalance between exploitation and exploration process, which would affect the optimization process[52]. In literature, [53] examined the hybridization of LSSVM with AFSA. Using electrical load data set as an empirical data, the produced results indicated that AFSA-LSSVM is capable to yield smaller average RE as compared to PSO-LSSVM.

On the other hand, ACO [49] imitates the foraging behavior of ants, which acts as agents in searching the shortest path between the colony and the food source. This algorithm is reported to

guarantee the convergence; however the time to converge is inconsistent [54]. Besides, as this algorithm is developed based on routing problems, the theoretical is reported to be quite difficult to be applied in other applications. In optimizing LSSVM hyper-parameters, [55] demonstrated the ACO which has been applied for share price prediction. As to eliminate the fluctuant component in raw data, Wavelet Transform (WT) was employed which served as preprocessor to LSSVM. The proposed model is measured based on MSE and the comparison was made against single SVM and ACO-SVM without WT. The results obtained indicated that ACO-LSSVM offered promising results by producing lower MSE.

Meanwhile, DE [50] is inspired by the mechanism of natural selection which considered as extension of GA. The difference between DE and GA is, in DE, all possible solutions have an equal chance in evaluation task, while in GA, the chance of updating the solution is depends on fitness value. The application of DE involved the predefined of several control parameters which affect the optimization process of DE and is revealed to have tendency of falling into local minimum [56]. This situation leads to the modification of DE (termed as CDEK) which has been demonstrated for optimizing the LSSVM hyper-parameters [56]. In the study, CDEK-LSSVM was tested on estimation of thermal process. Empirical experimentation suggested that the proposed model demonstrate better predictive power than the standard DE-LSSVM.

Previously, in 2004, the hybridization of DE-LSSVM has been presented which experimented on three benchmark functions [57]. Empirical results showed that the DE-LSSVM offered good convergence performance, however, in terms of prediction accuracy, the results obtained is unsatisfying.

Progressing further, [58] investigated the usefulness of online LSSVM in prediction of turbine heat rate. In the study, the GSA[59] was utilized to optimize the hyper-parameters of LSSVM. GSA is a population-based algorithm which is motivated from Newtonian gravity. Apart from population size and number of iteration, several control parameters need to be predefined by the user, which includes gravitational constant and inertia mass. Upon completing the experiment, the experimental results suggested that GSA-LSSVM offered better generalization capability as compared to the other identified prediction techniques.

In related work, [41] presented a hybrid of LSSVM with Bacterial Colony Chemotaxis (BCC) for short term load prediction. Categorized as bionic algorithm, BCC is colony-intelligent optimization which is based on bacterium's reaction to chemo attractants to find the optimum. Upon completing the experiment, lower MAPE was produced by BCC-LSSVM as compared to several identified techniques, which includes BPNN and CV-LSSVM. In terms of training time, the demerits of CV that was highlighted in [25] was proven as ineffective in the study where it took much longer time to complete the training task as compared to time recorded using BCC-LSSVM.

Meanwhile, in [60], the LSSVM is hybridized with Fruit Fly Optimization (FFO)[61] for electric load prediction. The FFO is inspired from the food searching behavior. In the study, the FFO is employed as an optimizer to LSSVM and 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 MAPE, MSE and Average Absolute Error (AAE).

As discussed above, both GA and PSO are widely utilized in optimizing the LSSVM hyper parameters. However, despite its extensive use of GA in literature, several deficiency of this technique has been reported. There is high possibility of GA to face with stagnation in the local minimum, premature convergence, low convergence speed and difficulty in parameters selection[53].As for PSO, it is also revealed that the PSO tends to fall into local minimum[62]. Besides GA and PSO, the application of other EC algorithms which includes AFSA, ACO, DE and GSA are also presented in literature. However, the applications of these techniques are less obtrusive except for Artificial Bee Colony (ABC)[63]. A study on the prediction of non-renewable commodity such as crude oil and gold has been reported in 2011using the ABC-LSSVM [64]. Results showed that the ABC-LSSVM possess higher accuracy than existing prediction models that includes Backpropagation Neural Network (BPNN) and Differential Evolution – Least Squares Support Vector Machines (DE-LSSVM).The work was later extended by the researchers to improve the search mechanism of ABC and this was presented as eABC-LSSVM [65,66]. The extension includes two sub algorithms termed as *lv*ABC-LSSVM [67] and *cm*ABC-LSSVM. The *lv*ABC algorithm was introduced to overcome the local optima problem by enriching the searching behaviour using Levy mutation. On the other hand, the *cm*ABC algorithm that incorporates conventional mutation addresses the over-fitting or under-fitting problem faced by ABC. Experiments on energy crude oil daily prices showed that *e*ABC-LSSVM produced promising MAPE, which is 6.3310%. With that, the PA obtained is 93.6690%. This is followed by ABC-LSSVM with 7.1769% of MAPE. The summary of literature based on EC algorithms is presented in Table 2.

Table 2: Time Series Prediction Technique using LSSVM Optimized by EC Algorithm

Authors	Research Area	Independent Variable	Method	Data Frequency	Evaluation Metric
Mustafa, et al. (2011)	Transmission loss prediction	Load curve and load demand	GA-LSSVM	Hourly	MSE
Mustafa, et al. (2012)	Reactive power prediction	Load curve and load demand	GA-LSSVM	Hourly	MSE
Yu, et al. (2009)	Stock market prediction	S&P 500 index, Dow Jones Industrial Average (DJIA) index and New York Stock Exchange (NYSE) index	GA-LSSVM	Monthly	Hit ratio
Fu, et al. (2010)	Ship motion prediction	Heading angle	AGA-LSSVM	Seconds	MAE
Sun & Zhang (2008)	Power spot prices prediction	Power spot prices	GA-LSSVM	Daily	MAE and MAPE
Xie, et al. (2009)	Dissolved gas concentration	Dissolved gas analysis data	GA-LSSVM	-	Absolute error

	prediction				
Yang, et al. (2010)	Conductivity and tensile strength prediction modeling for carbon fiber composites	Temperature, average fiber, fiber content, conductivity, and several others	GA-LSSVM	-	LMSE and GMSE
Wang, et al. (2005)	Chaotic time series data prediction	Henon map time series data	GA-LSSVM	-	RMSE
Liao &Balzen (2013)	Network flow prediction	Network access flow	Self-Adaptive GA-LSSVM	Hourly	RMSE and MAPE
Jiang & Zhang(2013)	Critical heat flux prediction	Mass flow rate, pressure, the heated length to tube diameter ratio and equilibrium quality	PSO-LSSVM	-	Relative Error
Xiang & Jiang (2009)	Water quality prediction	Water figures of Lixiu River	PSO-LSSVM	-	MAPE
Shen, et al. (2009)	Stock return prediction	Stock return of 5 days backward and closing prices of 3 days forward, and several others	WPOS-LSSVM	Daily	Error
Liao, et al.(2011)	Dissolved gases prediction	Dissolved gas data110kV and 220kV	PSO-LSSVM	-	MAPE and R^2
Liu & Yao (2009)	Chaotic time series prediction	Mackey-Glass time series data and Lorenz time series data	PSO-LSSVM	-	RMSE
Zhou & Shi (2010)	Ship motion prediction	A series of roll and pitch time series data of a real ship during her seakeeping trial	HPSO-LSSVM	-	RMSE

44

Zhu & Wei (2013)	Carbon price prediction	Carbon price data sets	PSO-ARIMA-LSSVM	Daily	RMSE, Dstat
Chen, et al. (2008)	Electricity load prediction	Electricity load data	AFSA-LSSVM	Hourly	RE
Fang & Bai (2009)	Share price prediction	Share price data	ACO-LSSVM	Daily	MSE
Dos Santos, et al. (2012)	Thermal process prediction	Temperature, relative humidity	CDEK-LSSVM	Second	R^2
Feoktistov&Janaqi (2004)	Prediction	Benchmark functions	DE-LSSVM	-	-
Zhang, et al. (2013)	Turbine heat rate prediction	Heat rate	GSA-LSSVM	Hourly	R^2, APE and MAPE
Li (2009)	Short-term load prediction	Load value of last three hours, the load value of the very hour in last three days, and several others	BCC-LSSVM	Hourly	MAPE and computational time
Li, et al. (2012)	Electric load prediction	Electricity consumption	FFO-LSSVM	Annually	MAPE, MSE and AAE
Mustaffa, et al. (2013)	Crude oil price prediction	Daily spot price and the variance	eABC-LSSVM	Daily	MAPE,MSE

3. CONCLUSION

As the success of LSSVM relies on the tuning of its parameters, optimizing its hyper-parameters, which are the regularization parameter, γ and kernel parameter, σ^2 is important. It is learned from existing work that optimizing the hyper-parameter is best implemented using the evolutionary algorithms. And, this includes the employment of swarm intelligence algorithms such as the Particle Swarm Optimization, Fish School Algorithm, Fruit Fly Optimization and Artificial Bee Colony.

REFERENCES

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

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

[3] H.-L. Xie, N. Li, F.-C. Lu, and Q. Xie, "Application of LS-SVM by GA for Dissolved Gas Concentration Forecasting in Power Transformer Oil," in Proceedings of the Asia-Pacific Power and Energy Engineering Conference (APPEEC). , 2009, pp. 1-4.

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

[5]   D. C. Montgomery, C. L. Jennings, and M. Kulahci, Itroduction to Time Series Analysis and Forecasting. New Jersey: John, Wiley &Sons, 2008.

[6]   S. Ismail, A. Shabri, and R. Samsudin, "A Hybrid Model of Self-Organizing Maps (SOM) and Least Squares Support Vector Machines (LSSVM) for Time Series Forecasting," Expert Systems with Applications, vol. 38, pp. 10574-10578, 2011.

[7]   A. Khashman and N. I. Nwulu, "Intelligent Prediction of Crudei Oil Price Using Support Vector Machines," in Proceedings of the 9th IEEE International Symposium on Applied Machine Intelligenc abd Informatics, Smolenice, Slovakia, 2011, pp. 165-169.

[8]   H. A. Khazem, "Using artificial neural networks to forecast the futures prices of crude oil," D.B.A. 3295968, Nova Southeastern University, United States -- Florida, 2008.

[9]   A. Alizadeh, M. Moghaddam, M. Khakzad, and V. Ebrahimipour, "A Flexible Neural Network-Fuzzy Mathematical Programming Algorithm for Improvement of Oil Price Estimation and Forecasting," Computers and Industrial Engineering, vol. 62, pp. 421-430, 2012.

[10] Y. Bao, X. Zhang, L. Yu, K. K. Lai, and S. Wang, "An Intergrated Model Using Wavelet Decomposition and Least Squares Support Vector Machines for Monthly Crude Oil Prices Forecasting," New Mathematics and Natural Computation, vol. 7, pp. 299-311, 2011.

[11] M. E. Malliaris and S. G. Malliaris, "Forecasting Inter-Related Energy Product Prices," The European Journal of Finance, vol. 14, pp. 453-468, 2008.

[12] J. Wang, T. Li, and R. Ren, "A real time IDSs based on artificial Bee Colony-support vector machine algorithm," in Proceedings of the Third International Workshop on Advanced Computational Intelligence (IWACI), 2010, pp. 91-96.

[13] I. Haidar and R. C. Wolff, "Forecasting of Crude Oil Price (Revisited)," in Proceedings of the 30th USAEE Conference, Washington D.C, 2011.

[14] L. Liu and W. Wang, "Exchange Rates Forecasting with Least Squares Support Vector Machine," in Proceedings of the International Conference on Computer Science and Software Engineering, 2008, pp. 1017-1019.

[15] C.-S. Lin, S.-H. Chiu, and T.-Y. Lin, "Empirical Mode Decomposition-based Least Squares Support Vector Regression for Foreign Exchange Rate Forecasting," Economic Modelling, vol. 29, pp. 2583-2590, 2012.

[16] A. Mellit, A. Massi Pavan, and M. Benghanem, "Least Squares Support Vector Machine for Short-Term Prediction of Meteorological Time Series," Theor Appl Climatol, vol. 2013, pp. 297-307, 2013.

[17] J. Wu and D. Niu, "Short-Term Power Load Forecasting Using Least Squares Support Vector Machines(LS-SVM)," in Proceedings of the Second International Workshop on Computer Science and Engineering (WCSE) 2009, pp. 246-250.

[18] X. Wang and H. Li, "One-Month Ahead Prediction of Wind Speed and Output Power Based on EMD and LSSVM," in Proceedings of the International Conference on Energy and Environment Technology (ICEET) 2009, pp. 439-442.

[19] D.-c. Wang, C.-x. Wang, Y.-H. Xie, and T.-Y. Zhu, "Air Temperature Prediction Based on EMD and LSSVM," in Proceedings of the Fourth International Conference on Genetic and Evolutionary Computing, 2010, pp. 177-180.

[20] M. T. Gencoglu and M. Uyar, "Prediction of Flashover Voltage of Insulators using Least Squares Support Vector Machines," Expert Systems with Applications, vol. 36, pp. 10789-10798, 2009.

[21] T. Van Gestel, J. A. K. Suykens, D. E. Baestaens, A. Lambrechts, G. Lanckriet, B. Vandaele, B. De Moor, and J. Vandewalle, "Financial time series prediction using least squares support vector machines within the evidence framework," IEEE Transactions on Neural Networks, vol. 12, pp. 809-821, 2001.

[22] A. S. Ahmad, M. Y. Hassan, and M. S. Majid, "Application of Hybrid GMDH and Least Squares Support Vector Machines in Energy Consumption Forecasting," in IEEE International Conference on Power and Energy (PECon), Kota Kinabalu, Sabah, Malaysia, 2012.

[23] Y. Wang and Y. Li, "Applying LS-SVM to Predict Primary Energy Consumption," in Proceedings of the International Conference on E-Product E-Service and E-Entertainment (ICEEE), 2010, pp. 1-4.

[24] N. Wu, X. Liang-fa, and W. Ji-lin, "The Application of Particle Swarm Optimization-Based RBF Neural Network in Fault Diagnosis of Power Transformer," in Proceedings of the 2nd IEEE

International Conference on Computer Science and Information Technology (ICCSIT), 2009, pp. 534-536.

[25] A. Lendasse, Y. Ji, N. Reyhani, and M. Verleysen, "LS-SVM Hyperparameter Selection with a Nonparametric Noise Estimator," in Proceedings of the 15th International Conference on Artificial Neural Networks: Formal Models and Their Applications (ICANN) Warsaw, Poland, 2005.

[26] L. Yu, H. Chen, S. Wang, and K. K. Lai, "Evolving Least Squares Support Vector Machines for Stock Market Trend Mining," IEEE Transactions on Evolutionary Computation, vol. 13, pp. 87-102, 2009.

[27] D. Karaboga, B. Gorkemli, C. Ozturk, and N. Karaboga, "A comprehensive survey: artificial bee colony (ABC) algorithm and applications," Artificial Intelligence Review, pp. 1-37, 2012/03/01 2012.

[28] D. Karaboga and B. Akay, "A comparative study of Artificial Bee Colony," Applied Mathematics and Computation, vol. 214, pp. 108-132, 2009.

[29] Z. Yang, X. S. Gu, X. Y. Liang, and L. C. Ling, "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, 2010.

[30] R. L. Haupt and S. E. Haupt, Practical Genetic Algorithms, Second ed. New Jersey: A John Wiley & Sons, Inc., Publication, 2004.

[31] M. W. Mustafa, M. H. Sulaiman, S. N. Abd. Khalid, H. Shareef, S. R. Abdul Rahim, and O. Aliman, "An Application of Genetic Algorithm and Least Squares Support Vector Machine for Tracing the Transmission Loss in Deregulated Power System," in Proceedings of the 5th International Power Engineering and Optimization Conference (PEOCO), Shah Alam, Selangor, 2011.

[32] M. W. Mustafa, M. H. Sulaiman, H. Shareef, and S. N. Abd. Khalid, "Reactive power tracing in pool-based power system utilising the hybrid genetic algorithm and least squares support vector machine," IET, Generation, Transmission & Distribution, vol. 6, pp. 133-141, 2012.

[33] H. Fu, S. Liu, and F. Sun, "Ship Motion Prediction Based on AGA-LSSVM," in Proceedings of the International Conference on Mechatronics and Automation (ICMA), 2010, pp. 202-206.

[34] W. Sun and J. Zhang, "Forecasting Day Ahead Spot Electricity Prices Based on GASVM," in Proceedings of the International Conference on Internet Computing in Science and Engineering (ICICSE), 2008, pp. 73-78.

[35] M. Zou, "Condition Prediction of Hydroturbine Generating Units using Least Squares Support Vector Regression with Genetic Algorithm," in Proceedings of the Eightth International Conference on Machine Learning and Cybernetics, Baoding, China, 2009.

[36] X. Wang, H. Zhang, C. Zhang, X. Cai, J. Wang, and J. Wang, "Prediction of Chaotic Time Series using LS-SVM with Automatic Parameter Selection," in Proceedings of the Sixth International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT), 2005.

[37] W. Liao and Z. Balzen, "LSSVM Network Flow Prediction Based on the Self-adaptive Genetic Algorithm Optimization," Journal of Networks, vol. 8, pp. 507-512, 2013.

[38] Kennedy and R. Eberhart, "Particle Swarm Optimization," in Proceedings of the IEEE International Conference on Neural Networks (ICNN), Australia, 1995, pp. pp. 1942-1948.

[39] Q. Chen, Y. Wu, and X. Chen, "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, 2008, pp. 66-70.

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

[41] Y. Li, "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), 2009, pp. 1-5.

[42] W. Shen, Y. Zhang, and X. Ma, "Stock Return Forecast with LS-SVM and Particle Swarm Optimization," in Proceedings of the International Conference on Business Intelligence and Financial Engineering, 2009.

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

[44] P. Liu and J. Yao, "Application of least square support vector machine based on particle swarm optimization to chaotic time series prediction," in Proceedings of the IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS). , 2009, pp. 458-462.

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

[46] B. Zhu and Y. Wei, "Carbon Price Forecasting with a Novel Hybrid ARIMA and Least Squares Support Vector Machines Methodology," Omega, vol. 41, pp. 517-524, 2013.

[47] L. X. Li, Z. J. Shao, and J. X. Qian, "An optimizing method based on autonomous animals: fish-swarm algorithm," System Engineering Theory and Practice, vol. 22, pp. 32-38, 2002.

[48] X. Li, Z. Shao, and J. Qian, "An Optimizing Method base on Autonomous Animates: Fish Swarm Algorithm.," Systems Engineering Theory and Practice, vol. 22, pp. 32-38, 2002.

[49] M. Dorigo and T. Stutzle, Ant Colony Optimization. Cambridge: MIT Press 2004.

[50] R. Storn and K. Price, "Differential Evolution- a simple and efficient heuristic for global optimization over continuous spaces," Journal of Global Optimization, vol. 11, pp. 341-359, 1997.

[51] E. Rashedi, H. Nezamabadi-pour, and S. Saryazdi, "GSA: A Gravitational Search Algorithm," Information Sciences, vol. 179, pp. 2232-2248, 2009.

[52] Y. Peng, "An Improved Artificial Fish Swarm Algorithm for Optimal Operation of Cascade Reservoirs," Journal of Computers, vol. 6, pp. 740-746, 2011.

[53] X. Chen, J. Wang, D. Sun, and J. Liang, "Time Series Forecasting Based on Novel Support Vector Machine Using Artificial Fish Swarm Algorithm," in Proceedings of the Fourth International Conference on Natural Computation (ICNC). , 2008, pp. 206-211.

[54] V. Selvi and R. Umarani, "Comparative Analysisi of Ant Colony and Particle Swarm Optimization Techniques," International Journal of Computer Applications, vol. 5, pp. 1-5, 2010.

[55] X.-Y. Fang and T. Bai, "Share Price Prediction using Wavelet Transform and Ant Colony Optimization for Parameters Optimization in SVM," in Proceedings of the Global Congress on Intelligent System, 2009.

[56] G. S. Dos Santos, L. G. J. Luvizotto, V. C. Mariani, and L. Dos Santos Coelho, "Least Squares Support Vector Machines with tuning based on Differential Evolution Approach Applied to the Identification of a Thermal Process," Expert Systems with Applications, vol. 39, pp. 4805-4812, 2012.

[57] V. Feoktistov and S. Janaqi, "Hybridization od Differential Evolution with Least Squares Support Vector Machines," in Proceedings of the Annual Machine Learning Conference of Belgium and The Netherlands (BENELEARN), Belgium, 2004, pp. 53-57.

[58] W. Zhang, P. Niu, G. Li, and P. Li, "Forecasting of Turbine Heat Rate with Online Least Squares Support Vector Machine based on Gravitational Search Algorithm " Knowledge-Based Systems, vol. 39, pp. 34-44, 2013.

[59] E. Rashedi, H. Nezahabadi-pour, and S. Saryazdi, "GSA: A Gravitational Search Algorithm," Information Sciences, vol. 179, pp. 2232-2248, 2009.

[60] H. Li, S. Guo, H. Zhao, C. Su, and B. Wang, "Annual Electric Load Forecasting by a Least Squares Support Vector Machines with a Fruit Fly Optimization Algorithm," Energies, vol. 2012, pp. 4430-4445, 2012.

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

[62] J.-B. Park, Y.-W. Jeong, J.-R. Shin, and K. Y. Lee, "An Improved Particle Swarm Optimization for Nonconvex Economic Dispatch Problems," IEEE Transactions on Power Systems, vol. 25, pp. 156-165, 2010.

[63] D. Karaboga and B. Basturk, "A Powerful and Efficient Algorithm for Numerical Function Optimization:Artificial Bee Colony (ABC) Algorithm," Journal Global Optimization, vol. 39, pp. 459-471, 2007.

[64] Z. Mustaffa and Y. Yusof, "Optimizing LSSVM using ABC for Non-Volatile Financial Prediction," Australian Journal of Basic and Applied Sciences vol. 5, pp. 549-556, 2011.

[65] Y. Yusof, S. S. Kamaruddin, H. Husni, Ku-Mahamud, Ku Ruhana, , and Z. Mustaffa, "Enhanced ABC-LSSVM for Energy Fuel Price Prediction," Journal of Information and Communication Technology., vol. 12, pp. 73-101, 2013.

[66] Z. Mustaffa, Y. Yusof, and S. S. Kamaruddin, "Enhanced Artificial Bee Colony for Training Least Squares Support Vector Machines in Commodity Price Forecasting," Journal of Computational Science, vol. 5, pp. 196-205, 2014.

[67] Z. Mustaffa and Y. Yusof, "Levy Mutation in Artificial Bee Colony Algorithm for Gasoline Price Prediction," in Proceedings of the Knowledge Management International Conference (KMICe), Johor Bahru, Johor, Malaysia, 2012.