SUPPORT VECTOR REGRESSION WITH ARTIFICIAL BEE COLONY FOR SUPPLY CHAIN DEMAND FORECASTING

Krinal Patel	**Prof. Risha Tiwari**
M.E Student	Assistant Professor
Department of Computer Engineering	Department of Computer Engineering
Hasmukh Goswami College of Engineering	Hasmukh Goswami College of Engineering
Gujarat Technological University	Gujarat Technological University
Ahmadabad, Gujarat	Ahmadabad, Gujarat

ABSTRACT

In this paper, we introduce the machine learning technique of time series forecasting Support Vector Regression (SVR). There are no general guidelines are available to choose the free parameters of an SVR model. We use the Artificial Bee Colony (ABC) algorithm to optimize the SVR parameters. Artificial Bee Colony (ABC) algorithm which inspired from the behaviour of honey bees swarm is presented. ABC is a stochastic population-based evolutionary algorithm for problem solving. ABC selects best free parameters for SVR to avoid over-fitting and local minima problems and improve prediction accuracy. Experimental results show that the proposed model SVR-ABC outperforms the existing model SVR-PSO.

KEYWORDS— Forecasting, Supply Chain Management, Support Vector Regression, Artificial Bee Colony

I.   INTRODUCTION

Demand forecasts play a crucial role in supply chain management. Nowadays, there are different techniques for calculating forecasts. ARIMA models defined by Box and Jenkins (1994 [8]), can be used for this purpose. But, with these traditional methods, the construction of supply chain demand forecasting (SCDF) model may be difficult due to its non-linear, dynamic and complicated characteristics. ARIMA method deals with linear, but it is difficult to deal with nonlinear feature in time series. Thus, methods based on artificial intelligence techniques like artificial neural network (Minsky and Papert in 1969 [8][3]), genetic algorithms (Holland in 1975 [4]) and support vector machine (Vapnik et a1. 1997 [6]) are used to improve the performance of forecasting in non linear context. The most used method is artificial neural network (ANN). Neural network as a nonlinear system model has good self-learning, adaptive ability and generalization, fault-tolerant ability, which can better explain the complicated nonlinear relationship between variables and the influence factors. However, Hu and Zhang (2008[1]) showed that ANN has inherent drawbacks, such as local optimization solution, lack generalization, and uncontrolled convergence. In most cases ANNs suffer from over-fitting problem due to the large number of parameters to fix, and the little prior user knowledge about the relevance of the inputs in the analysed problem.

The recent machine learning technique, support vector machine (SVM), which overcomes the drawbacks of neural networks. Support vector machine (SVM) is a novel neural network algorithm based on statistical learning theory (Vapnik, 2000)[6]. Support vector machine (SVM) is suitable for solving the problems of small samples, nonlinear, high dimension, and local minimum. In the modelling of financial time series using SVR, one of the key problems is how to select model parameters correctly, which plays an important role in good generalization performance.

To overcomes the drawback in determining the parameters of SVR, this study optimization technique. Today there are number of optimization techniques in market as Particle swarm optimization (PSO), Evolutionary Algorithm , Ant colony optimization , Artificial Bee Colony, Bat algorithm, Gravitational search algorithm, Bird Flocking, Animal Herding, Bacteria Growth, Fish Schooling etc. Artificial Bee

Colony is very simple, flexible, robust optimization algorithm. It also has some other features like use of fewer control parameters as compared to other optimization algorithms, ease of hybridization with other optimization algorithms, ability to handle the objective cost with stochastic nature and ease of implementation with basic mathematical and logical operations.

II. THE HYBRID MACHINE LEARNING TECHNIQUE
A. SUPPORT VECTOR MACHINE

The support vector machine (SVM) is a recent tool from the artificial intelligence field which use statistical learning theory that has been successfully applied to many fields and it recently of increasing interests of researchers: It has been introduced by Vapnik et a1.(1992) and was first applied to pattern recognition (classification) problems, recent research has yielded extensions to regression problems, including time series forecasting[1].

SVMs belong to the general category of kernel methods. A kernel method is an algorithm that depends on the data only through dot-products. When this is the case, the dot product can be replaced by a kernel function which computes a dot product in some possibly high dimensional feature space. SVMs maximize the margin around the separating hyper plane. SVM is high accuracy, ability to deal with high-dimensional data.

B. SUPPORT VECTOR REGRESSION

A version of a SVM for regression has been proposed in 1997 by Vapnik, Steven Golowich, and Alex Smola. This method is called support vector regression (SVR). The idea of SVR is based on the computation of a linear regression function in a high dimensional feature space where the input data are mapped via a nonlinear function   where $X_{i,}$ is the ith input vector and $Y_{i,}$ is the ith desired value. The objective is to find an unknown regression function, y = g(x), The SVR function is shown as follows

$$f(x) = \omega \emptyset(x) + b$$

Where f(x) denotes the forecasting values Φ(x) is the features, which is nonlinear mapped from the input space x. b is a constant, and w denotes the weight vector estimated by minimizing the regularized risk function [2]:

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

Subject To:
$$y_i - (<\omega, \phi(x_i)> +b) \leq \varepsilon + \xi_i$$
$$(<\omega, \phi(x_i)> +b) - y_i \leq \varepsilon + \xi_i^*$$
$$\xi_i \xi_i^* \geq 0$$

Where C and ε are user defined parameters The parameter ε is the difference between actual value and value calculated from regression function. The constant C determine the trade-off between the flatness of f and amount up to which deviation larger than ε are tolerated ξ_i denotes the training error above ε whereas ξ_i^* denotes the training error below $-\varepsilon$ and   n represent the number of samples. SVR avoids under fitting and over fitting of the training data by minimizing the regularization term $\frac{1}{2} ||\omega||^2$ as well as the training error $C \sum_{i=1}^{N}(\xi_i + \xi_i^*)$.

After the quadratic optimization problem with inequality constraints is solved, the parameter vector ω is obtained by [1]:

$$\omega = \sum_{i=1}^{n}(\beta_i^* - \beta_i)\emptyset(x_i)$$

Where β_i^* and β_i are obtained by solving a quadratic program and are the Lagrangian multipliers finally, the SVR regression function is obtained as the following equation in the dual space:

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

Where $K(x_i, x_j)$ is called the kernel function: The value of the kernel equals the inner product of two vectors x_i and x_j in the feature space $\emptyset(x_i)$ and $\emptyset(x_j)$. There are four type of kernel function. We use RBF (Radial basic function).

$$K(x_i,x) = \exp\left(-\gamma|x_i - x_j|^2\right) \quad (RBF)$$

Where γ is the adjustable parameter?

C. ARTIFICIAL BEE COLONY

The Artificial Bee Colony (ABC) algorithm is a stochastic, population-based evolutionary method proposed by Karaboga in the year 2005. ABC algorithm is simple and very flexible when compared to other swarm based algorithms. This method has become very popular and is widely used, because of its good convergence properties. Artificial Bee Colony (ABC) is a swarm intelligence based algorithms inspired by the intelligent behavior of honey bees. Honey bees use several mechanisms like waggle dance to optimally locate food sources and to search new ones. Waggle dance is a means of communication among bees by which the successful foragers share the information not only about the direction and distance of the food sources but also about the amount of nectar available to the other foragers. This information exchange among bees helps them in detecting the optimal food locations.ABC has the advantages of very simple, strong robustness, fast convergence, high flexibility and fewer setting parameters[12].

In ABC three artificial agents are defined named as the employed, the onlooker and the scouts. A bee waiting on the dance area for making a decision to choose a food source is called onlooker and one going to the food source visited by it before is named employed bee. The other kind of bee is scout bee that carries out random search for discovering new sources [9].

In general the position of ith food source is represented as Si = (Si1, Si2, …. ,SiD). Information is shared by the employed bees after returning to the hive, onlooker bees go to the region of food source explored by employed bees at Si based on probability Pi defined as [12]

$$Pi = \frac{\text{fit}_i}{\sum_{k=1}^{FS} \text{fit}_i}$$

Where FS is total is total number of Food Sources. Fitness value fit_i is calculated by using the equation [12]

$$\text{Fit}_i = \frac{1}{1 + f(S_i)}$$

Where $f(S_i)$ denotes the objective function considered. The onlooker finds its food source in the region of Si by using the following equation [12]

$$S_{new} = S_{ij} + r^* \left(S_{ij} - S_{kj} \right)$$

Where Snew is the new food source exploited by onlooker and k is the solution in the neighbourhood of i, r* is the random number in the range -1 to +1 and j is the dimension of the problem considered.

III.LITERATURE SURVEY

Related work for this research includes demand forecasting of supply chain in general. Particularly, works which used SVR for demand forecasting.

Malek and Afia (2014) [1] introduce the machine learning technique of time series forecasting Support Vector Regression (SVR) and use the Particle Swarm Optimization (PSO) algorithm to optimize the SVR parameters. The experimental data should be necessary divided into the two subsets: the training data and the testing data. The forecasting accuracy is measured by the mean absolute percentage error (MAPE) on the testing data. The paper takes the retail sales data published by the Census Bureau. These data are available from the official website of the bureau. SVR-PSO gives better results than SVR-GA for this sample and SVR-PSO gives better minimum values than the SVR-GA method.

Horng and Tsung (2011) [2] proposed a time series forecasting approach by integrating particle swarm optimization (PSO) and support vector regression (SVR). SVR has been widely applied in time series predictions. However, no general guidelines are available to choose the free parameters of an SVR model. The proposed approach uses PSO to search the optimal parameters for model selections in the hope of improving the performance of SVR. The prediction performance is evaluated using the following statistical metrics, namely, the root mean square error (RMSE), mean absolute difference (MAD), mean absolute percentage error (MAPE), directional symmetry (DS), correct up trend (CU) and correct down trend (CD). This study compared the forecasting results of the proposed method with those of the traditional SVR method using prediction error and prediction accuracy as criteria. Compared to the

traditional SVR model, experimental results showed that the proposed PSO-SVR model can produce lower prediction error and higher prediction accuracy in the datasets. Besides, the CPU time for the proposed model is about eight times faster than that when using the traditional SVR method.

CHEN and Liu (2013) [3] combines the support vector regression machine (support vector regression, SVR) and Particle Swarm Optimization algorithm, (Particle Swarm Optimization, PSO) to propose PSO-SVR coal logistics demand forecasting model which is suitable for the learning of small samples.SVR model and then use the particle swarm algorithm to optimize model parameters. The results show that PSO-SVR is superior to BP in forecast accuracy and error. In a word, PSO-SVR is a very valuable coal logistics demand forecasting model with its faster training speed and higher efficiency.

Cai, Lu and Zhang (2009) [4] proposed Support vector regression optimized by genetic algorithm (G-SVR) to forecast tourism demand. This study examines the feasibility of SVR in tourism demand forecasting by comparing it with back-propagation neural networks (BPNN).The G-SVR model provides lower forecasting error than the ANN model. The superior performance of G-SVR model over ANN approaches is mainly due to the following causes: SVR models have nonlinear mapping capabilities and can easily capture atmospheric corrosion depth data patterns compared to other models.

YanGao (2009) [5] proposed Support vector regression optimized by genetic algorithm (G-SVR) to forecast freight volume and adopt genetic algorithm(GA) to seek the optimal parameters of SVR in order to improve the efficiency of prediction. In the study, we normalize the experimental data to improve generalization capability of SVR. GA is adopted to seek the optimal parameters of SVR in order to improve the efficiency of prediction. In the freight volume forecasting, the experimental results show that G-SVR can achieve greater forecasting accuracy than artificial neural network, grey model.

Nasimul Hasan,Nayan,Risul (2015) [6] author took only the data of March to October of each year.80% of the data were considered as training data and the rest 20% as testing data. The result of the experiment done with data (2008-2014) of Chittagong, Bangladesh show that the proposed model can forecast more accurately in comparison with regular technique used.

Gao and Feng (2009) [7] present a modelling and forecasting method of urban logistics demand based on regression SVM is presented. The SVM network structure for forecasting of urban logistics is established. Moreover, authors propose a self-adaptive parameter adjust iterative algorithm to confirm SVM parameters, thereby enhancing the convergence rate and the forecasting accuracy. we can see the mean squared errors of training samples of SVM is bigger than the RBNN, but to the mean squared errors of testing samples, the RBFNN is bigger. It shows that the generalized ability of the regression SVM is

Karin Kandananond (2012) [8] present artificial neural network (ANN) and support vector machine (SVM), and a traditional approach, the autoregressive integrated moving average (ARIMA) model, were utilized to predict the demand for consumer products. The SVM outperformed the other two methods in almost every category of product.

Kuldeep, Sunita and Kapil Sharma (2015) [9] introduce Artificial Bee Colony optimization algorithm.ABC is well suited for general assignment problem, cluster analysis, constrained problem optimization, structural optimization, and advisory system.

Osman ,Omar,Mustafa (2014) [10] proposed LSSVM and optimize parameter by PSO and ABC. After that LSSVM is optimized by ABC algorithm to be used in the prediction of daily stock prices. Artificial Bee Colony algorithm (ABC), a population-based iterative global optimization algorithm is used to optimize LSSVM for stock price prediction. ABC algorithm is used in selection of LSSVM free parameters C (cost penalty), ϵ (insensitive-loss function) and γ (kernel parameter). The proposed LSSVM-ABC model Convergence to a global minimum can be expected. Also proposed model overcome the over-fitting problem which found in ANN, especially in case of fluctuations in stock sector. LSSVM-ABC algorithm parameters can be tuned easily. Optimum found by the proposed model is better than LSSVM-PSO and LSSVM. Proposed model converges to global minimum faster than LSSVM-PSO model.

IV. SVR-ABC FOR SCDF

Step 1.To determine the input and output variables

Choosing Watson Analytics sales product data and take demand, price and elasticity value as an input and predict demand, price and elasticity value as an output.

*S*tep 2.Selection of the learning samples.

Selecting the appropriate learning sample is bedrock of the SVR model. The learning samples are divided into training samples and testing samples. The front is used for training and the latter used for testing the validity of the model. This paper selects the Watson Analytics sales product data as the test sample.

Step 3. The choices of kernel function

The choices of kernel function have greatly accuracy influence on the SVR model prediction. The radial basis kernel function shows better performance than other kernels in dealing with nonlinear samples. So this paper chooses radial basis function (RBF) kernel as kernel function of this prediction model.

Step 4.To Select the structural parameters:

The main structure parameters of the SVR prediction model is the gamma parameter γ. So, selecting gamma parameter this paper uses Artificial Bee Colony (ABC) method to determine the structure parameters.

Step 5.Model training

Through the training sample data, the paper use radial function as the kernel and put the software package of LIBSVM into Net beans. According to the structure parameters optimized by ABC in Step 4, the model is trained. Then, put the training model into Step 6 and training results into Step 7.

Step 6.Test prediction.

Use SVR model trained by Step 5 to predict the test samples and put the prediction results to Step 7.

Step 7.The training accuracy test

In order to evaluate the training and testing effects, the paper use the MAPE to test the results from Step 6 and Step 7. If the results do not meet the requirements of precision, jump to Step 4 to optimize structure parameters, for retraining and the second prediction. If meeting the accuracy, go to the Step 8.

Step 8.The output results of training and testing results.

This paper uses the relative error to express the predictive validity .The comparison between SVR-ABC AND SVR-PSO shown in figure 1.

V. RESULT AND ANALYSIS

We can see from the Figure.1 and Table I that the SVR-ABC gives better minimum values than SVR-PSO method.

In base paper, SVR-PSO is better than SVR-GA and we prove that SVR-ABC is better than SVR-PSO. So, We can say that SVR-ABC is better than SVR-GA.

Table I Comparison Of SVR-ABC and SVR-PSO Results

	SVR-ABC	SVR-PSO
MAPE	0.002235	0.484916

Fig: 1 Comparison of execution time of SVR-ABC and SVR-PSO
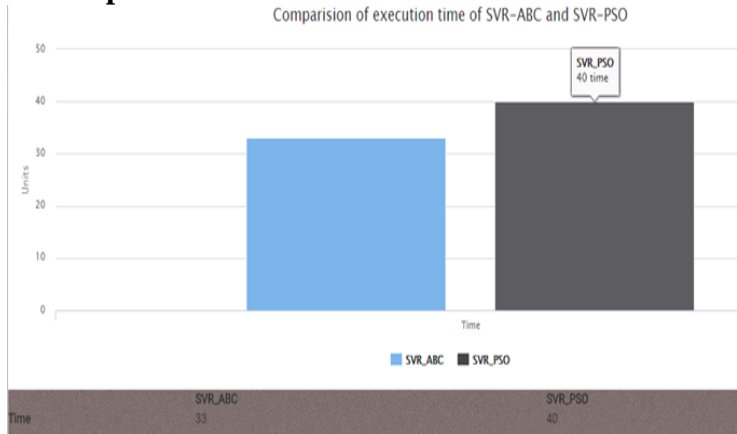
VI. CONCLUSION

In this paper, we find hybrid machine learning technique SVR-ABC which implement Support Vector Regression which parameters optimized by Artificial Bee Colony. Watson Analytics sales product data is

used in this study for evaluating performance of proposed method. We compare SVR-ABC with SVR-PSO model. We can conclude that Artificial Bee Colony is the best algorithm to choosing the γ (Gamma) parameter of SVR model. The experimental result shows that SVR-ABC model outperforms   SVR-PSO model.

VII.   FUTURE WORK

Use Support Vector Regression method which parameters optimized by different optimization algorithm and improve the result of demand forecasting and also use other forecasting methods which improve the accuracy of demand forecasting.

REFERENCES

1. Malek Sarhani and Abdellatif El Afia "Intelligent System Based Support Vector Regression For Supply Chain Demand Forecasting" 978-1-4799-4647 -1/14/IEEE Year 2014
2. Horng-I Hsieh and Tsung-Pei Lee "A Hybrid Particle Swarm Optimization and Support Vector Regression Model for Financial Time Series Forecasting" Vol. 2, No. 2; ISSN 1923-4007 , May 2011
3. CHEN Pei-you and Liu Lu   "Study on Coal Logistics Demand Forecast Based on PSO-SVR" 978-1-4673-4843-0/13/$31.00/ IEEE Year 2013
4. Zhong-jian Cai,Sheng Lu Xiao-bin Zhang "Tourism demand forecasting by support vector regression and genetic algorithm" 978-1-4244-4520-2/09/$25.00/ IEEE Year 2009
5. YanGao "Forecasting of freight volume based on support vector regression optimized by genetic algorithm" 978-1-4244-4520-2/09/$25.00/ IEEE Year 2009
6. Nasimul Hasan, Nayan Chandra Nath, Risul Islam Rasel "A Support Vector Regression Model for Forecasting Rainfall" 978-1-4673-9257-0/15/$31.00/ IEEE Year 2015
7. Meijuan Gao and Qian Feng "Modeling and Forecasting of Urban Logistics Demand Based on Support Vector Machine" 978-0-7695-3543-2/09 $25.00/IEEE Year 2009
8. Karin Kandananond "A Comparison of Various Forecasting Methods for Auto correlated Time Series" Volume. 4, Year 2012
9. Kuldeep Singh,Sunita Chaudhary and Kapil Sharma "Application of Artificial Bee Colony Optimization Technique: Survey" IEEE Year 2015
10. Osman Hegazy, Omar S. Soliman and Mustafa Abdul Salam " LSSVM-ABC Algorithm for stock price prediction" International Journal of Computer Trends and Technology (IJCTT) – volume 7 number 2– Jan 2014
11. Kamalam Balasubramani "A Comprehensive review of Artificial Bee Colony Algorithm" ISSN 2277-3061 Volume 5, No. 1, May -June, 2013
12. S. Oshaba, E. S. Ali and S. M. Abd Elazim "Artificial Bee Colony Algorithm Based Maximum Power Point Tracking in Photovoltaic System" ISSN: 2224-350X Volume 10, Year 2015