International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 02, 2021

Predictive Analysis Model for Oil price Forecasting using a Hybrid Model

Anitha Kumari K[1], Hinduswari M[2]

[1]Associate Professor, Dept. of IT, PSG college of Technology, Coimbatore, TN, India
[2]Software Engineer, Netlink digital solution, Chennai, TN, India

Email: [1]kak.it@psgtech.ac.in, [2]hindhu.psgtech@gmail.com

Abstract— Petroleum plays a vital role in the economy of our country. It is one of the important sources of the energy and act as vital raw material for a number of industries. Every economic sector in the world is dependent on crude oil, any increase or decrease in the price of crude oil has a ripple effect on the global economy. Hence accurately predicting the price petroleum become the hot issues in many countries around the world. To grasp the trend of petroleum price and reduce the negative impact of petroleum price changes. In this paper a new improved hybrid model is proposed by combining Particle Swarm Optimization (PSO) and Back Propagation (BP)-Artificial Neural Network(ANN) techniques to predict the oil price.

Keywords— Predictive analysis, oil price forecasting, improved Hybrid model.

1. INTRODUCTION

Petroleum is one of the indispensable energy for development of world economy and politics. It is also used for a wide range of applications, most-notably powering internal combustion engine. The International Energy Agency(IEA) estimates that oil will supply 30 percent of the world's energy mix in 2030, reports National Research Council. Around two-thirds of oil in the U.S. and Canada is used for transportation. Oil is mostly used for power generation and space heating in other countries. Moreover, oil is a valuable product for the agriculture industry, which provides food for people across the globe. Hence, to forecast the petroleum price in an efficient manner is essential.

Forecasting oil price for both long and short term is possible with specific degree accuracy. Therefore, research on the prediction model with higher precision, better performance, efficiency, correlation and wider adaptability are still the first-line goal in academia and industrial community. The price of oil changes constantly**,** depending on how much is produced and consumed. Many factors can influence the price of oil. Especially wars and conflicts in oil-producing countries can lead to a rise in oil prices. In this paper, to predict the future price of petrol, a new improved hybrid model is proposed by combining Artificial Neural Network(ANN) and Particle Swarm Optimization(PSO) algorithm. For high performance this prediction model is proposed based on PSO and ANN techniques with the assistance from the petroleum industry.

RELATED WORK

In this section, we present a brief review of the related and recent studies.

Mohamed El Hédi Arouri [1] pointed out that accounting for the potential instabilities and structural breaks in the unconditional variance of oil spot and future returns almost leads to out-of-sample forecast gains, as compared to the benchmark GARCH(1,1) which provides less accurate volatility forecasts whatever the forecasting horizon considered. Ardalan Tebyanian and Fares Hedayati [3] ensemble methods work better in tuning parameters of regression models in predicting the price of crude oil than a simple cross validation. However, the computational cost is high.

Wang et al [4] present a hybrid methodology to forecast crude oil monthly prices. The model consists of a combination of three separate components, Web mining from which the authors extract rule based system, in addition ANN, and ARIMA models. These three components work disjointedly, and then integrated together to get the final results.

Adnan Khashman and Nnamdi I. Nwulu [4] obtained an overall correct prediction rate of 81% is considered as successful taking into account the nonlinearity of the problem. The authors claim the lead-lag4 relation could only hold for no more than a half hour, and their results confirm that changes in futures prices help predict the changes in the spot price.

Lu Xue-tong [5] forecasting oil price a hybrid model by combining Backpropagation and standard particle swarm optimization is proposed. The standard particle swarm optimization has been improved by self-adaptive weight update strategy and it is used to improve the convergence speed. Conventional algorithm results in low performance due to premature convergence.

Siddhivinayak Kulkarni and Imad Haidar [6] proposed an ANN model for short-term crude oil price prediction. In addition, crude oil future prices contain information about the spot price direction in the short-term.

John Wei-Shan H, Yi-Chung Hu and Ricky Ray-Wen Lin [7]

The focus of the work is to apply neural networks for predicting crude oil futures prices. In this Recurrent Neural Network(RNNs), the predictive power improves by increasing the training time. A possible explanation for this phenomenon is the existence of a large difference between the predicted value and the actual value.

I. Background Techniques

A. *Particle Swarm Optimization*

Particle swarm optimization algorithm comprises of a collection of particle revolves around in the search space influenced by its gbest and pbest value. For each and every particle the velocity is assigned to it and it is represented as $v_i = v_{i1}, v_{i2}, .v_{in}$ where particle is represented as $x_i = x_{i1}, x_{i2} .... x_{in}$. Each particle has local memory (pbest) and global memory (gbest). Initially the pbest value is calculated randomly and based on the local memory value, the gbest to be calculated. The velocity is calculated in the range of $\left[v_{\max}, v_{\min}\right]$ and it is estimated using below equation. $v_i = v_i + \varphi_1 \times rand \times (pbest -$ (1) $2 \times rand \times$ $(gbest - x_i)$ $x_{it} + 1 = x_{it} + v_{it} + 1$

Convergence to local optimum is analyzed and modification is required to get the local optimal solution due to premature convergence problem. In the second form of PSO three factor to be considered they are local optimum solution and global optimum solution and its geometric center of the swarm. The swarm size is kep (1) every particle in the search area and it is moved to a new area to ascertain acceleration coefficient between every particle.

a) Improved PSO Algorithm: The standard PSO algorithm has its own limitation of premature convergence problem when it is applied to the higher dimension problem. The global optimal solution is not improved by increasing the iteration value. Hence, primary PSO is modified slightly to improve the performance of the algorithm for calculating gbest and pbest

International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 02, 2021

value from which the geometric center is calculated. Further each particle will be selected by predefined probability from the population and it is added to the velocity of the particle. Then the maximum number of neighborhood is calculated and it's divided by least value that fit for the algorithm. For each neighborhood the pbest is calculated.

The improved PSO is as follows :

Step 1: Initialize population and position

Step 2: Calculate the fitness value.

Step 3: Compare the fitness value with its pbest value.

Step 4: For each particle compare the current best position of pbest with whole swarm current gbest value.

Step 5: Adjust the speed of position according to the formula

$$v[k+1] = w.v[k] + c.rand().(pbest[k] - present[k] +$$
$$c2.rand().(gbest[k] - present[k])$$
$$present[k+1] = present[k] + v[k+1] \qquad (2)$$

V [] represent the speed of the particle. Present [] represent the position of the current particle, rand () represents a random number (0, 1) and c1 and c2 are learning factor.

Step 6: The weight is to be calculated and particle swarm node position vector is changed with k_{th} iteration

Step 7: If stopping criteria is not met, go to step2. Otherwise, stop the algorithm.

B. *Artificial Neural Network*

The mathematical model for artificial neural network is to calculate the weighted sum of the input neuron based on the transfer function. To assign a transfer function sigmoid function is used. ANN algorithm is capable of learning, but need to be trained. One positive aspect is, there is no need to adjust the base parameter. The topology of the artificial neural network is of two types: feed forward neural network and Backpropagation algorithm.

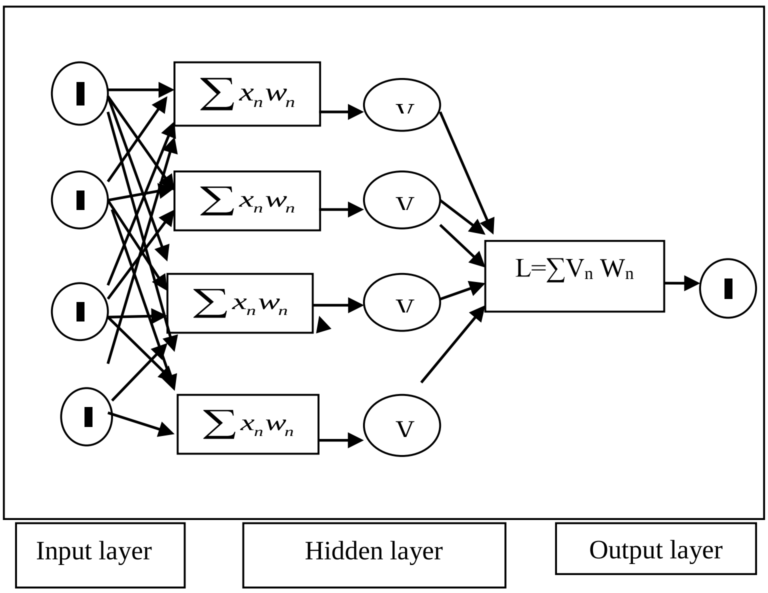

Fig. 1. Architecture of neural network

The input layer has six attributes passed through the weighted sum function which is in hidden layer, the synaptic weight and bias value is added to the obtained optimum result because error value is calculated at the end of output layer function. A detailed architecture is shown in figure 1. From the recent studies, it has been proven that the ANN provides several advantages such as extracting the relevant information from the dataset and for preprocessing. It is an efficient tool that could most suitable for pattern recognition, forecasting and linear regression model. By taking the advantage of both particle swarm optimization and artificial neural network algorithm, a hybrid model is proposed in this paper to predict the petrol price. This algorithm is capable of separating and assigning the huge data into testing and training dataset. Finally, it is expected to attain a better accuracy with less computational cost comparatively with other non-standard model for oil price forecasting techniques.

2. PROPOSED METHODOLOGY

In this paper, to predict the petrol price a hybrid model is proposed by combining particle swarm optimization and artificial neural network technique. The proposed algorithm takes the advantage of both of the techniques for predicting the oil price. In artificial neural network the important feature to get a better accuracy is transfer function: synaptic weight and bias value. To set the function it has to codified based on the requirement of the application. Setting the transfer function as per the need of application is the complex task in neural network. According to the survey, particle swarm optimization is the best algorithm for adjusting the parameter. Each neuron is directly related to the dimension of the individuals and size of individual depends on input pattern. It revolves around the search area and finds the best solution. It is expected that ANN produce the better accuracy. For better understanding, block diagram is presented in figure 2.

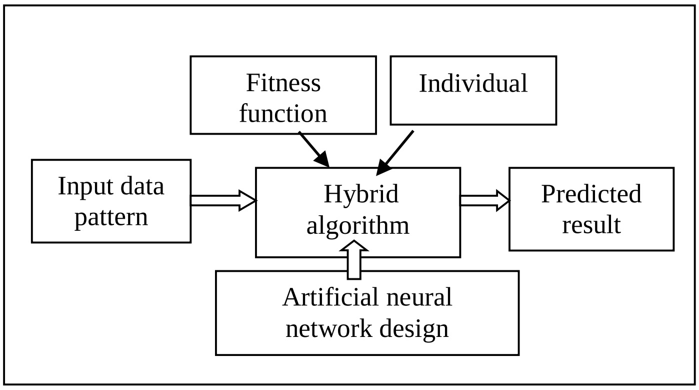

Fig. 2. Overall block diagram

A comprehensive PSO algorithm is presented below:

Algorithm:
Step1. Initialize inertial factor *w*, learning factor 1*c*, 2 *c*, The population scale *P*, the maximum iteration time, particle speed and position.
Step2. The particle in the swarm is mapped as network weight and threshold vector of BP neural network.
Step3. Input training sample to train the network, calculate fitness value of each particle.
Step4. For each particle, if it fitness value is better than *pbest*, *pbest* is position of the particle at present. For the whole particle swarm, if there is one particle with better fitness value than *gbest*, *gbest* is position of the particle at present.
Step5. If it meets the error precise, the global optimal particle is mapped as initial weight and threshold of neural network.
Otherwise update particle speed and position to generate the new swarm, then check whether the new swarm has cluster tendency.
If there is cluster tendency, turn to step 3, Otherwise instigate to process position mutation.

3. EXPERIMENT AND RESULT DISCUSSION

The experiment consists of historical dataset from *Western Texas* intermediate. This paper adopts three layer structure of ANN. According to the specific oil price forecasting problems, we choose The Organization of the Petroleum Exporting Countries (OPEC) crude oil supply, crude oil consumption, The Organisation for Economic Co-operation and Development (OECD) oil supply, crude and consumption of OECD, U.S. crude oil consumption and U.S. crude oil inventories as input dataset. The input of the neuron is assigned to the hidden layer. The transfer function used is sigmoid function. The sigmoid function is the mathematical function especially suitable for logistic function as shown in figure 3. Output of output layer is WTI crude oil price and the number of output layer nodes is 1. The number of hidden layer nodes is determined by means of heuristic optimization analysis.

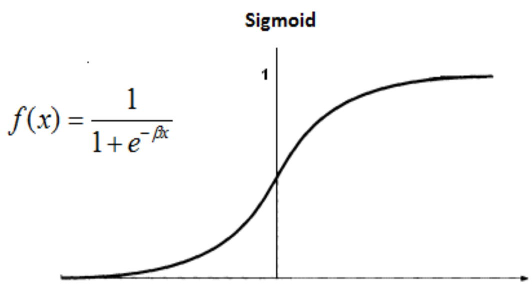

Fig.3. Sigmoid function

Since the artificial neural network has hidden layer the sigmoid transfer function is used. It is a mathematical function response and ratio between input and output. A feed-forward ANN model consists of multi-neurons in multilayer. The starting layer, called input layer, is connected to an output layer through a number of hidden layers. The number of hidden layers is determined by heuristic analysis. The resultant output as expected and actual value for each individual is shown in table1. The weighted sum is calculated and updated further. Each individual must be selected based on their fitness, then sum of squared error value is obtained for predicted result. Synaptic weight is to be updated and passes through the PSO algorithm to optimize the result set. Finally, training algorithm computes the error for prediction which is a minimized one as showcased in Table 1.

TABLE I. SQUARED ERROR VALUE

S. No	***Expected output***	***Actual output***	***Weighted value***
1	0.9999999	0.9761482	0.0168147
2	0.99907223	0.96623198	0.5034066
3	0.98123671	0.9701243	0.88633743
4	0.99702234	0.96907223	-0.50036332
5	0.98796712	0.96921504	-0.37780052
6	0.98623198	0.96907223	-0.25880964
7	0.98251567	0.96854464	-0.117123374
8	0.99562143	0.97623198	1.440205621
9	0.98623198	0.9601243	0.886276459
10	0.99123456	0.95329865	1.29728970

The actual value is obtained from the hybrid algorithm and is shown in figure 4. The transfer function for this function calculated using the weight of the neural network design. Initially, each value is selected based on the local memory value (pbest) at random in the range(0,6), later based on the obtained pbest value, the global optimum value is obtained at the end of last iteration.

The sum squared error value is predicted based on the number of epoch value. Epoch is the number of the iteration where the globally optimal solution is reached and it is calculated using formula

$$\frac{\sum_{i=1}^{d}(y_i - y_i')^2}{d} \qquad (3)$$

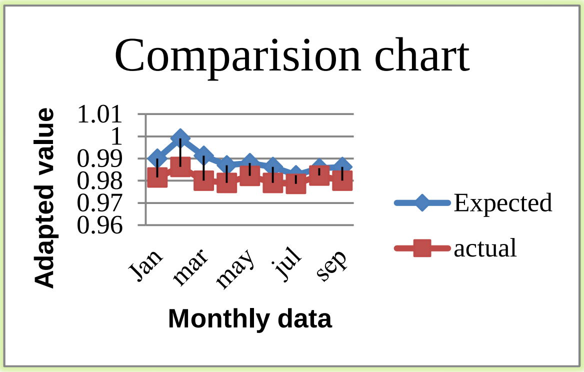

Fig. 4. Actual versus predicted value

Then the bias value is added to the output layer, where size of the output layer depends on the input pattern. Particle swarm optimization fitness function is defined in the particle process module in the algorithm produces the acceptable accuracy value for fitting the neural network algorithm. This hybrid model has less computational complexity where testing and training set data are used to train hybrid (PSO-ANN) algorithm. Global optimal particle is mapped as the initial weight and threshold of the neural network. The most widely used activation functions for ANN in the hidden layer are the sigmoid functions and the hyperbolic tangent. By adjusting the parameter of the fitness algorithm and activation function of network structure the proposed model provide acceptable efficiency. Also, the actual value is determined with less fluctuation. The monthly average value and number of epoch value is plotted as shown in figure 4.

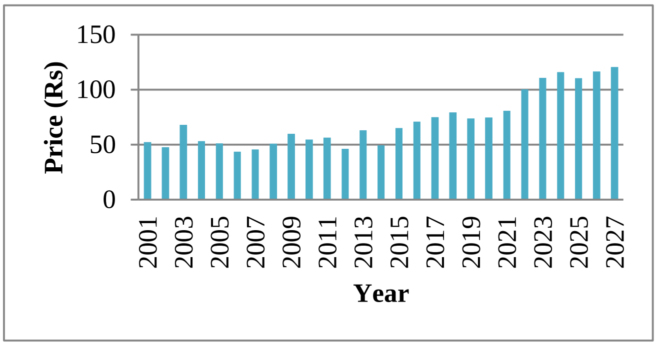

Fig. 5. Predicted result

The expected value to be predicted is based on the historical data that comprises of the parameter monthly price of crude oil, taxes and retail charges. The proposed model process the data values on the basis of network adoption layer and provide result as single value and it is shown in figure 5.

4. CONCLUSION

In this paper, a hybrid model is proposed by combining particle swarm optimization and artificial neural network algorithm to predict the oil price with better accuracy. Dataset is collected monthly wise from Western Texas Intermediate. The fitness function and activation function are codified for prediction and the result from the neural network is optimized using PSO. From the predicted value the graph generated and the squared sum error value is obtained.

The result shows that the acceptable accuracy value than the traditional work. Further for optimization the neural network value is allowed to map the fitness function and find the global optimal solution in the search area (gbest). Results shows that overall prediction result has improved in the proposed algorithm and the feasibility of the petroleum price prediction model is developed. Each individual is identified by the position and velocity of the particle in the given search space of the PSO algorithm. Also, it has direct communication with other individuals in improved artificial neural network algorithm. Comparing with conventional techniques, the proposed technique shows a better performance.

5. REFERENCES

[1] Beatriz A. Garro 1 and Roberto A. Vázquez 2 ,Design "Artificial neural network usin Particle Swarm optimization", 3rd ed., vol2 computational intelligence and neuroscience ,2015.
[2] Lu Xue-tong and Dong Wan-li, "Oil price Forecasting based on Particle Swarm Neural network," in Magnetism, vol. III, Seventh International Conference on Measuring Technology and Mechatronics Automation,2015.
[3] Darbellay, G. A. and M. Slama. "Forecasting the Short-term Demand for Electricity: Do Neural Networks Stand a Better Chance?" *International Journal of Forecasting,* Vol. 16, pp.71– 83.2012
[4] Niaz Bashiri Behmiri and José R. Pires Manso "Crude oil forcasting tecnique :a comphrehensive review of the literature",Research review-2014.
[5] Lean Yu a, Shouyang Wang a, Kin Keung Lai b."Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm" 2008 Elsevier
[6] Garro B., Sossa H., Vazquez R. Design of artificial neural networks using differential evolution algorithm. In: Wong K., Mendis B., Bouzerdoum A., editors. *Neural Information Processing. Models and Applications*. Vol. 6444. Berlin, Germany: Springer; 2010. pp. 201–208. (Lecture Notes in Computer Science).
[7] Garro B. A., Sossa H., Vazquez R. A. Design of artificial neural networks using a modified particle swarm optimization algorithm. Proceedings of the International Joint Conference on Neural Networks (IJCNN '09); June 2010; Atlanta, Ga, USA. IEEE; pp. 938–945.
[8] Eberhart R. C., Shi Y., Kennedy J. *Swarm Intelligence*. 1st. Boston, Mass, USA: Morgan Kaufmann; 2001. (The Morgan Kaufmann Series in Evolutionary Computation).
[9] Higashi N., Iba H. Particle swarm optimization with Gaussian mutation. Proceedings of the IEEE Swarm Intelligence Symposium (SIS '03); April 2003; IEEE; pp. 72–79.
[10] D. McNeils, "Neural networks in finance gaining predictive edge in the market," Elsevier Acadmic Press, Massachusetts, 2005.
[11] S. Wang, L. Yu, and K. Lai. "Crude oil price forecasting with TEI@I methodology." Journal of Systems Science and Complexity, vol. 18(2), 145-166, 2009
[12] S. Wang, L. Yu, and K. Lai. "Crude oil price forecasting with TEI@I methodology." Journal of Systems Science and Complexity, vol. 18(2), 145-166, 2007.
[13] Source http://www.opec.org/opec_web/en/data_graphs/333.htm
[14] Goel, P. and Thareja, R. (2017) "Analysis of Various Data Mining Techniques using Novel Ratings Prediction", IARS' International Research Journal. Vic. Australia, 7(2). doi: 10.51611/iars.irj.v7i2.2017.75.
[15] Chang, C.C. and C.J. Lin. (2005). "LIBSVM: a Library for Support Vector Machines." http://www.csie.ntu.edu.tw/~cjlin.

[16] Chang, C.C. and C.J. Lin. (2005). "LIBSVM: a Library for Support Vector Machines." http://www.csie.ntu.edu.tw/~cjlin.
[17] Li, J. B.,Liu, X. G.,Jiang, H. Q.,et al. Melt index prediction by adaptively aggregated RBF neural networks trained with novel ACO algorithm. Journal Of Applied Polymer Science, 2012,125(2): 943-951.
[18] The International Energy Agency, 2004, "Analysis of the impact of high oil prices on the global economy," Report
[19] Lubna A Gabralla1, Ajith Abraham1, 2," Computational Modeling of Crude Oil Price Forecasting: A Review of Two Decades of Research. International Journal of Computer Information Systems and Industrial Management Applications.ISSN 2150-7988 Volume 5 (2013), pp. 729-740
[20] Xie, W., Yu, L., Xu, S., & Wang, S. "A new method for crude oil price forecasting based on support vector machines." *In Computational Science–ICCS ,Springer Berlin Heidelberg,*pp. 444-451, 2006
[21] He,K., Yu, L.and Lai, K.K., "Crude oil price analysis and forecasting using wavelet decomposed ensemble model". *Energy*. PP.46, 564-574, 2012
[21] Lin, A. "Prediction of international crude oil futures price based on GM (1, 1)." In *Grey Systems and Intelligent Services, . GSIS . IEEE International Conference on*, pp. 692-696. IEEE, 2009.