Forecasting Crude Oil Prices Using Improved Deep Belief Network (IDBN) and Long-Term Short-Term Memory Network (LSTM)

Mohammad Mahdi Lotfi Heravi
Department of Management, Science and Technology
Amirkabir University of Technology (Tehran Polytechnic)
Tehran, Iran
Mahdi.lotfi@aut.ac.ir

Mahsa Khorrampanah
Department of Electrical Engineering
Imam Reza International University
Mashhad, Iran
Khorrampanah_m@yahoo.com

Monireh Houshmand
Department of Electrical Engineering
Imam Reza International University
Mashhad, Iran
m.hooshmand@imamreza.ac.ir

2022 30th International Conference on Electrical Engineering (ICEE) | 978-1-6654-8087-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICEE55646.2022.9827452

***Abstract*—Historically, energy resources are of strategic importance for the social welfare and economic growth. So, predicting crude oil price fluctuations is an important issue. Since crude oil price changes are affected by many risk factors in markets, this price shows more complicated nonlinear behavior and creates more risk levels for investors than in the past. We propose a new method of prediction of crude oil price to model nonlinear dynamics. The results of the experiments show that the superior performance of the model based on the proposed method against statistical previous works is statistically significant. In general, we found that the combination of the IDBN or LSTM model lowered the MSE value to 4.65, which is 0.81 lower than the related work (Chen et al. protocol), indicating an improvement in prediction accuracy.**

Keywords—Crude oil price forecast, Deep learning model, Improved Deep Belief Network (IDBN), Return Nervous Network (RNN), Long-Term Short Shor-Term Memory Network (LSTM)

I. INTRODUCTION

Machine learning is a group of artificial intelligence algorithms that provides the system with the benefits of automated learning of knowledge without [1]. Many in-depth learning techniques have been presented, with promising results for a variety of applications such as Visual Data Processing, Natural Language Processing (NLP), Speech & Audio Processing, and so on. Popular deep learning models include the Convolution Neural Networks (CNNs), Deep Belief Networks (DBNs), and Deep Recursive Neural Networks (DRNNs). Among them, Deep Belief Networks (DBNs) have shown satisfactory performance. This can be associated with the use of foot propagation in deeply mononuclear neural networks, such as long learning times, the need for a substantially labeled data set for training, and selective techniques. The Short-Term Permanent Memory Network (LSTM) solves some basic mathematical problems in long sequence modeling [3-6].

In [3], the study uses the deep learning model to describe nonlinear characteristics of oil price movements. Using the deep learning model, they suggest a new model for combining oil price. Experimental results show that the model improves the prediction. In [7], a new and modern method based on online media text mining method to predict the price of crude oil, is presented. In particular, this is the first study to apply in-depth learning algorithems to predict crude oil price and extract hidden patterns in online media of news using a CNN.

To predict crude oil prices, in [8] the authors propose a new approach that combines the complete decomposition of the experimental state with adaptive noise (CEEMDAN) and infinite incremental gradient (XGBOOST) which is called XGBOOST-CEEMDAN. In [9] the paper proposes a new multiscale method for estimating value at risk. This method uses the advantages of the variable state analysis model to extract and model the risk parameters in the multi-scale field, while the specific features of these risk parameters are modeled using GARCH-ARMA models. In [10], the long-term memory of the deep learning algorithm is used to predict the fluctuating behaviors of crude oil prices, and this is a novel way to predict price in energy markets using deep learning networks. By using LSTM and complex network analysis [11] proposes a hybrid model for forecasting crude oil price. In [12] the authors use in-depth learning techniques to obtain crude oil price behavior. They use long-term memory with comparison with mean motion, linear regression and mean motion of self-decreasing complex.

In this paper, we present a new combination learning method based on deep learning to model the dynamics of oil price. We integrate the deep learning model by considering the linear and nonlinear properties of the data records, and the predictive results. To evaluate performance results of the proposed model, experimental studies have been performed. The performance of the IDBN and Neural Network, is better than the performance of the Deep Belief Network.

The paper is organized in such a way that in the second part, we explain the proposed solution. In the third part, we evaluate the proposed algorithm, and finally conclusions and suggestions for future work are presented.

II. PROPOSED PROTOCOL

The proposed deep learning model includes the RBM, the Convolution Neural Networks (CNNs), the Deep Belief Networks (DBNs), the Deep Reversal Neural Networks (DRNNs), and the Deep Belief Networks (DBNs). This has been shown to be extremely effective. This can be attributed to backward propagation issues in Deep-Layer Neural Networks. Similarly, the LSTM can solve some of the basic mathematical problems in modeling long sequences. IDBN and LSTM are used to predict crude oil price to improve the forecast accuracy. These two algorithms will have a share of the final answer based on the same dataset. To determine the role of each algorithm (IDBN, LSTM) for increasing the accuracy of the system, we define W1 and W2 as following:

$$w_1 = \frac{MSE\ of\ IDBN}{MSE\ of\ IDBN + MSE\ of\ LSTM}$$

$$w_2 = \frac{MSE\ of\ LSTM}{MSE\ of\ IDBN + MSE\ of\ LSTM}$$

$$If\ w_1 > w_2 \rightarrow Y = w_2$$

$$If\ w_1 < w_2 \rightarrow Y = w_1$$

$$If\ w_1 = w_2 \rightarrow Y = w_2 or\ w_1$$

Due to the difficulty of deep network visualization in the network training step, training parameters have an important effect on the model. To improve DBN performance, we propose an adaptive semi-supervised method for training the RBM. So, we use the learning parameter in the training process of each RBM structure adaptively and we enter the dataset into the RBM with noise. We minimize errors between outputs from unsupervised training to optimize parameters. In traditional DBN, the parameters of the network are error propagation, which decreases with increasing depth and inability to prevent noise. We improve the efficiency of the forward stack training step as follows: First, we adjust the adaptive learning parameter based on the before and after reconstruction error. Then, we calculate the difference between the value of the partial derivative function (P (h | v)) in distribution $<\cdot>_{P(h|v)}$ and the value of the partial derivative function under the distribution of the reconstruction model ($<\cdot>_{recon}$) as following [14]:

$$\Omega = <\cdot>_{P(h|v)} - <\cdot>_{P(v|h)} \cdots \cdots$$

Second, to optimize the parameters, we enter the sample and the noisy version of the same sample into the RBM and minimize the error between the unsupervised training outputs. Predicting nonlinear components is done as follows: 1) In the normal time series, the AIC minimization evaluation is used to select the optimal lag order. 2) Trainings are determined, training data is used to teach the network parameters. 3) Crude oil price for the experimental set with deep learning network is trained and predicted using the rolling window method.

A. Dataset

In this research, historical data of WTI crude oil price market (from www.quandl.com) has been used. We divided the main data into two parts, in which 70% of the data was used as a training set and the remaining 30% of the data were used as testing data.

B. Simulation Setup

The implementation process was performed by MATLAB 2018b. For evaluating the performance of the proposed algorithm, we have chosen the parameters identical to the parameters of the article [3]. The number of hidden layers of the LSTM is chosen as 2, hidden neurons to 4, input nodes to 2, output nodes to 1, and every 47 nodes to 4, the Mean Squared Error Method (MSE) as a function for target value of the model is used. Table I shows the parameter settings (IDBN).

TABLE I. PARAMETER SETTINGS OF IDBN

Parameter	Primary value
Total member	100
The maximum itration	2000
Number of nodes	100, 50, 20, 10
alpha	0.1
decay	0.0001
k	1

C. Evaluation Criteria

1) Mean Squared Error (MSE): For a regression model, the value obtained based on the sum of the residual squares. The following relation is calculated for the dependent variable Y:

$$MSE_b = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y_l})^2 \quad (1)$$

2) Kurtosis: We use quadratic torque to calculate the Kurtosis. If μ is the mean and φ is the standard deviation of the random variable X, then the *Kurtosis* (Kur(X)) will be as follows (E refers to the mathematical expectation of a random variable):

$$Kur(X) = E\left[\left(\frac{X-\mu}{\varphi}\right)^4\right] = \frac{\mu_4}{\varphi^4} \quad (2)$$

3) Skewness: Skewness (statistics) indicates the degree of asymmetry of the probable distribution. If the data are symmetric with respect to the mean, the skewness will be zero. If μ is the mean and φ is the standard deviation of the random variable X, then the Skewness (Y_1) will be as follows:

$$Y_1 = E\left[\left(\frac{X-\mu}{\varphi}\right)^3\right] = \frac{\mu_3}{\varphi^3} \quad (3)$$

4) Standard Deviation: The standard deviation obtains the difference of each data from the mean. The mean value of these second powers is squares (variances) and its positive second root is called the standard deviation. If we denote the number of data by n and the value of each by X_i, the standard deviation formula is as next:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \hat{X})^2}{n}} \quad (4)$$

III. RESULTS

The results are shown in Table II. The p value for both JB and BDS tests is less than 0.05. This shows that the crude oil price does not follow the normal distribution and it depends on nonlinear dataset.

TABLE II. FEATURES OF COMPLEX NONLINEAR DYNAMIC DATA

Data	MSE	S.D.	S.	K.	pJB	pBDS
WTI	4.6446	0.0727	-2.0206	6.8477	0.001	0

The results in Table III indicate that the best performance is obtained by using the IDBN-LSTM model, which has the minimum MSE among all models. In general, the MSE of the DBN-based model is lower than that of the LSTM-based models. This leaks the superior predictive power of the DBN model compared to the LSTM model in the WTI crude oil market. The better performance of the model based on the proposed method against the article [3] is statistically significant. We generally find that combining the IDBN or LSTM model lowers the MSE value, which indicates an improvement in forecast accuracy. It should be mentioned, that the prediction accuracy of the base article method is less than the proposed combined method. The predicted values of the proposed model as well as the main efficiency are shown in Fig. 1. Also, Fig. 1 and Fig. 2 show that the IDBN-based model in the proposed method follows the actual efficiency closer than the LSTM model. The prediction model of the price of the oil based on the LSTM method (Fig. 2) generates predictions higher transient events with a significant range of dataset. In changing steady-state oil data, the LSTM model may not be adaptable enough to the steady-state changes to bring in new existing changes. The market is dominated by short-term memory behaviors. Our results from Fig. 1 and Fig. 2 are consistent with the results of the Table III. As shown in the Table III, the MSE of the proposed protocol is the least, so, the dispersion of the results in the prices of the oil dataset is low (Fig. 1). To compare with the LSTM protocol, the dispersion of the prediction of the price is low when we use the proposed protocol. Also, the variability of the prediction of the price is closer to the reality than LSTM.

TABLE III. COMPARISON OF THE PERFORMANCE OF DIFFERENT MODELS

Model	DBN	LSTM	Article [3]	Proposed protocol
MSE	5.3236	5.5219	5.4504	4.6446

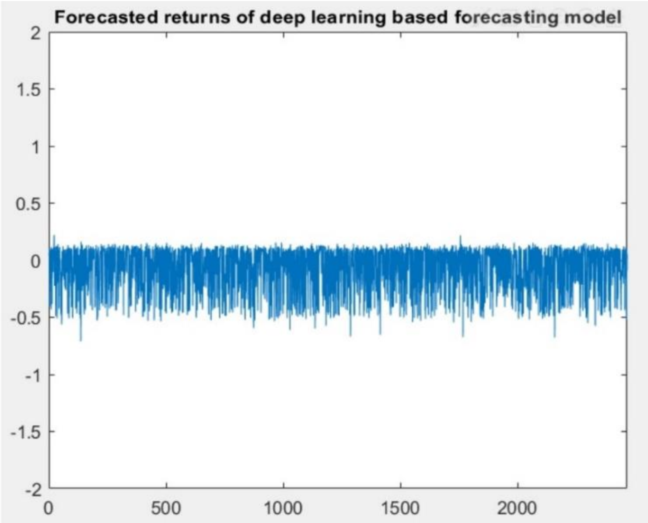

Fig. 1. Predicted returns of the proposed method

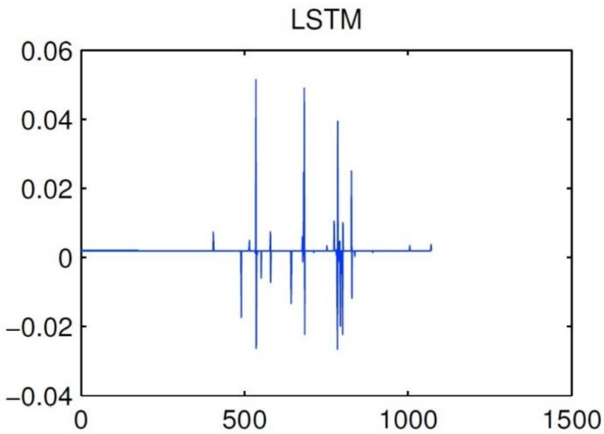

Fig. 2. Predicted returns of the LSTM method

IV. CONCLUSION AND FUTURE WORKS

In this paper, an emerging deep learning model is used to predict the price of oil. We use two specific deep learning models, namely deep belief network and return neural network which are useful in modeling nonlinear dynamics. A hybrid model is created that combines the predictions of deep learning models. We find that the deep learning model in crude oil price models leads to better results. We also understood that the functionality of the learning model is dependent to the parameters. Increasing the model complexity with more neurons and layers does not necessarily reach to more nonlinear modeling accuracy. This performance limitation may lead to a few number of types of deep learning models. The experimental results show that the desired performance is obtained using the IDBN-LSTM model, which has the minimum MSE among all models. In general, the MSE of the DBN-based model is lower than the LSTM-based models. This indicates the superior predictive power of the DBN model compared to the LSTM model in the WTI crude oil market. In comparison of the article [3], the performance of the proposed protocol is significantly high. We generally find that combining the IDBN or LSTM model lowers the MSE value, indicating an improvement in forecast accuracy.

Interestingly, the prediction accuracy of the basic article method is less than the proposed combined method.

In the future work, we intend to combine the proposed method with optimization algorithms and increase the accuracy of prediction. For this aim, we will select the best parameters for deep belief networks of optimization algorithms.

REFERENCES

[1] Sh. Dargan, M. Kumar, M. Rohit Ayyagari, and G. Kumar. "A survey of deep learning and its applications: A new paradigm to machine learning," Archives of Computational Methods in Engineering (2019): 1-22.

[2] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. Presa Reyes, Mei-Ling Shyu, Shu-Ching Chen, and S. S. Iyengar. "A survey on deep learning: Algorithms, techniques, and applications," ACM Computing Surveys (CSUR) 51, no. 5 (2018): 1-36.

[3] Y. Chen, K. He, and G. KF Tso. "Forecasting crude oil prices: a deep learning based model," Procedia computer science 122 (2017): 300-307.

[4] L. Yu, Sh. Wang, and K. K. Lai. "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm," Energy Economics 30, no. 5 (2008): 2623-2635.

[5] Y. Wei, Y. Wang, and D. Huang. "Forecasting crude oil market volatility: Further evidence using GARCH-class models," Energy Economics 32, no. 6 (2010): 1477-1484.

[6] I. Haidar, S. Kulkarni, and H. Pan. "Forecasting model for crude oil prices based on artificial neural networks," In 2008 International Conference on Intelligent Sensors, Sensor Networks and Information Processing, pp. 103-108. IEEE, 2008.

[7] X. Li, W. Shang, and Sh. Wang. "Text-based crude oil price forecasting: A deep learning approach," International Journal of Forecasting 35, no. 4 (2019): 1548-1560.

[8] Y. Zhou, T. Li, J. Shi, and Z. Qian. "A CEEMDAN and XGBOOST-based approach to forecast crude oil prices," Complexity 2019 (2019).

[9] Y. Zou, L. Yu, G. KF Tso, and K. He. "Risk forecasting in the crude oil market: A multiscale Convolutional Neural Network approach," Physica A: Statistical Mechanics and its Applications 541 (2020): 123360.

[10] Zh. Cen, and J. Wang. "Crude oil price prediction model with long short term memory deep learning based on prior knowledge data transfer," Energy 169 (2019): 160-171.

[11] M. Bristone, R. Prasad, and A. A. Abubakar. "CPPCNDL: Crude oil price prediction using complex network and deep learning algorithms," Petroleum (2019).

[12] D. Makala, and Z. Li. "ECONOMIC FORECASTING WITH DEEP LEARNING: CRUDE OIL," MATTER: International Journal of Science and Technology 5, no. 2 (2019).

[13] Y-X. Wu, Q.-B. Wu, and J.-Q. Zhu. "Improved EEMD-based crude oil price forecasting using LSTM networks," Physica A: Statistical Mechanics and its Applications 516 (2019): 114-124.

[14] T. Yang, and Sh. Huang. "Fault Diagnosis Based on Improved Deep Belief Network," In 2017 5th International Conference on Enterprise Systems (ES), pp. 305-310. IEEE, 2017.