Crude Oil Price Prediction Using LSTM Networks

Varun Gupta, Ankit Pandey

***Abstract*—**Crude oil market is an immensely complex and dynamic environment and thus the task of predicting changes in such an environment becomes challenging with regards to its accuracy. A number of approaches have been adopted to take on that challenge and machine learning has been at the core in many of them. There are plenty of examples of algorithms based on machine learning yielding satisfactory results for such type of prediction. In this paper, we have tried to predict crude oil prices using Long Short-Term Memory (LSTM) based recurrent neural networks. We have tried to experiment with different types of models using different epochs, lookbacks and other tuning methods. The results obtained are promising and presented a reasonably accurate prediction for the price of crude oil in near future.

***Keywords*—**Crude oil price prediction, deep learning, LSTM, recurrent neural networks.

I. Introduction

PREDICTION of crude oil prices has been a wide topic for ages. People use their intuition and lot of techniques to guess the prices of crude oil. It takes a lot of knowledge about the crude oil to accurately predict it. Predicting the crude oil price is very significant in various economic, political and industrial areas, both for crude oil importer and exporter countries [1]. Since crude oil is the most important strategic resource around the globe; it has become the crucial commodity for the world's economy. Thus, prediction of prices of crude oil has always been considered as a very exciting and challenging task which drew the curiosity of professionals, researchers and organizations all over the world [2]. Moreover, crude oil volatility has a critical impact on macroeconomic parameters such as such as inflation, unemployment, exchange rate, economic growth of countries whose economy rely heavily on crude oil export or import. Thus, crude oil price prediction can help governments of countries of the world in economic policymaking and make quick and operative economic decisions to hedge against probable risk in these economic parameters [3]. Therefore, forecasting of crude oil prices is quite useful and is also the objective of this paper.

In this paper, we have used LSTM [4]-[6] based recurrent neural networks for the purpose of crude oil price prediction. Recurrent neural networks (RNN) have been proved to be one of the most powerful models for processing time-series based sequential data. LSTM is one of the most successful RNN architectures. LSTM introduces the memory cell, a unit of computation that replaces traditional artificial neurons in the hidden layer of the network. With these memory cells, networks are able to effectively associate memories and input remote in time, hence suit to grasp the structure of data dynamically over time with high prediction capacity.

The paper is further organised into the following sections. Section II discusses the related works done in this field and provides background for the proposed network. Section III introduces the proposed architecture of the network based on LSTM. The results obtained from the study are discussed in Section IV and Section V concludes the paper and discusses the future scope of the work.

II. Related Work and Background

A. Related Work

Crude oil forecasting is an important topic in financial and economic studies. Many studies have been performed to forecast the prices of crude oil. After performing several tests, in 2005, Moshiri and Foroutan [7] concluded that future prices time series is stochastic and non-linear. They compared ARMA and GARCH techniques to ANN and found that ANN performed better for crude oil price forecasting.

Kulkarni and Haidar [8] presented a model for forecasting crude oil spot price direction in the short-term, up to three days ahead based on multilayered feedforward neural network. They tested the relation between crude oil future prices and spot price. They found the evidence that future prices of crude oil contain new information about oil spot price detection.

Hamdi and Aloui [9] performed a literature survey of numerous studies done on forecasting of crude oil price using artificial neural networks (ANN) until 2014. From the survey, they concluded that crude oil market is the most volatile commodity and forecasting oil price using nonlinear models such as ANN is the most suitable choice.

Abdullah and Zeng [10] proposed Hierarchical Conceptual (HC) and Artificial Neural Networks-Quantitative (ANN-Q) model based on machine learning and computational intelligence techniques to predict the monthly WTI crude oil price for every barrel in USD. The results obtained from their study validated the effectiveness of data selection process by the proposed model which successfully extracts a comprehensive list of key factors that cause the crude oil price market to be volatile.

Chen et al. [11] proposed a crude oil price forecasting model based on the deep learning model. They were able to analyze and model the crude oil price movement using the proposed deep learning model. They used the proposed model to capture the unknown complex nonlinear characteristics of the crude oil price movements. They evaluated the

Varun Gupta is with the Chandigarh College of Engineering and Technology (Degree Wing), Sector 26, Chandigarh 160019 India (corresponding author, phone: 91-9646047091; e-mail: varun3dec@gmail.com).

Ankit Pandey is with the Chandigarh College of Engineering and Technology (Degree Wing), Sector 26, Chandigarh 160019 India (e-mail: ankpandey@outlook.com).

performance of the proposed model using the price data in the WTI crude oil markets.

In recent times, LSTM based RNN have been widely used in sequence based problems such as machine translation [12], text classification [13], power demand forecasting [14], stock market prediction [15], question-answering [16], music notes recognition [17], behavior recognition of robots [18] and sentiment analysis [19]. In this paper, LSTM based RNN have been used for crude oil price forecasting. These networks are explained in the rest of this section.

B. Recurrent Neural Networks

RNN are different from feedforward networks. They use their internal memory to predict things. They are very good at tasks at which humans are not good at such as handwriting recognition and speech recognition. They were initially developed in 1980 [21]. These networks make use of sequential information available to them. Traditionally, we assumed that inputs do not depend on each other. But that was not a valid assumption. As if we want to predict the next words in a sentence we must know the previous words. They can be thought of having a memory which stores the information for future use. Fig. 1 shows a simple RNN model.

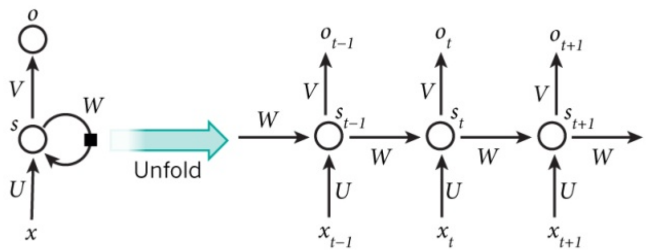

Fig. 1 RNN [6]

There exist various extensions of RNN. One of them is the Bidirectional RNN. In these networks, output at time t may depend on future inputs as well. Fig. 2 depicts the basic bidirectional RNN model.

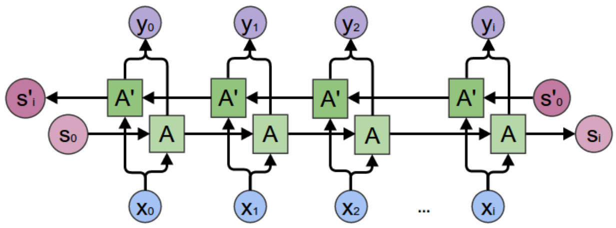

Fig. 2 Bidirectional RNN [6]

The other popular variant of the RNN is the deep RNN. In these recurrent networks, there exist multiple layers per time step. Fig. 3 shows a model of a deep RNN.

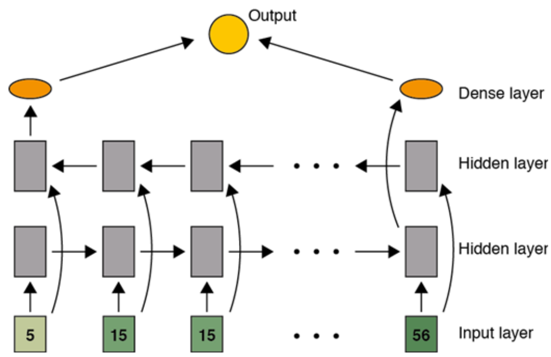

Fig. 3 Deep RNN

C. LSTM Networks

The most popular and widely used type of RNN is the LSTM and these types of recurrent networks have been used for this study. These networks learn order dependence in sequence prediction problem. The LSTM networks are able to solve two major issues encountered in RNN i.e. vanishing gradients and exploding gradients. The key to the solution of these problems was the internal structure that has been used in LSTM. Fig 4 shows a simple LSTM network. In this, there exists one input layer, one hidden layer and one output layer. This most simple architecture of LSTM networks is known as vanilla LSTM which performs very well in all sequence related prediction problems.

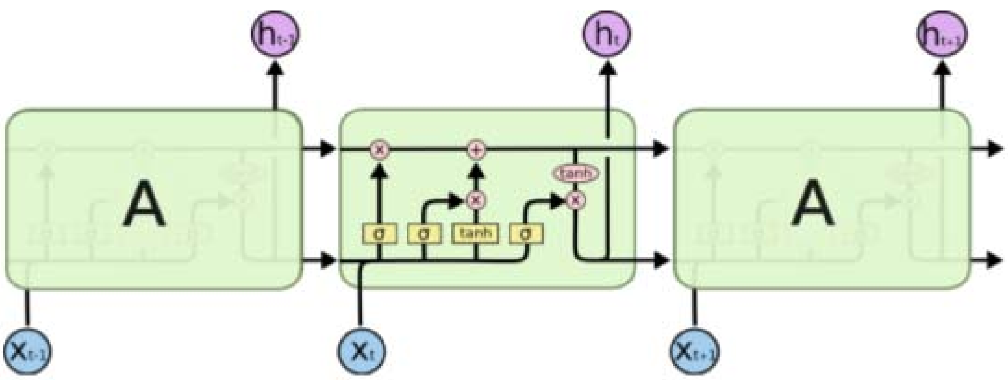

Fig. 4 LSTM [6]

III. Proposed Architecture

In this paper, LSTM based architecture has been used for prediction of crude oil price movements. The proposed architecture consists of four layers of LSTM layers followed by a dense layer with ten neurons and at the end dense layer with only one neuron. The proposed architecture is shown in Fig. 5.

All the inputs to the proposed network were normalized to achieve the best results. A dataset for crude oil prediction was obtained from [20]. The sample of the dataset is shown in Fig. 6.

The crude closed price column extracted is shown in Fig. 7. The crude oil dataset is depicted graphically in Fig. 8.

IV. Results and Discussions

Before deciding the final architecture of the network, a number of different configurations of the network were tested. Initially, four LSTM layers were used with lookback of 10 and 100 epochs. The training score obtained was 224.19 RMSE and the testing score was 550.50 RMSE. The obtained results are shown in Fig. 9. Then, the proposed network was experimented by using 6 LSTM layers and with 20 lookback and 100 epochs. The result of experimentation was that training score obtained was 235.12RMSE and testing score obtained was 793.24RMSE. These results showed that with an increase in lookback, accuracy of the network actually decreased. The obtained results are shown in Fig. 10.

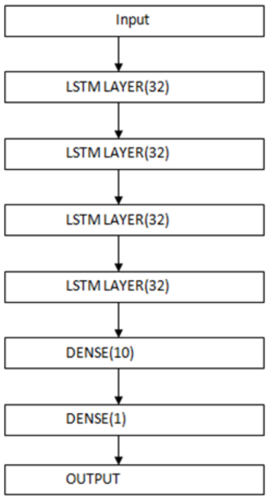

Fig. 5 Proposed Architecture

Date	Nifty_Close_Price	Nifty_Return	Gold_Close_Price	Gold_Return	Crude_Close_Price	Crude_Return
2006-04-12 00:00:00	3380	-0.0379967	8815	0.036624	3115	0.0429645
2006-04-19 00:00:00	3535.85	0.045078	9486	0.0733622	3342	0.0703403
2006-04-26 00:00:00	3555.75	0.00561229	9502	0.00168528	3282	-0.0181164
2006-05-03 00:00:00	3634.25	0.0218368	9806	0.0314921	3283	0.000304646
2006-05-10 00:00:00	3754.25	0.0324858	10336	0.0526385	3231	-0.015966
2006-05-17 00:00:00	3635.1	-0.0322519	10335	-9.67539e-05	3205	-0.0080796
2006-05-24 00:00:00	3115.55	-0.154231	9520	-0.0821413	3220	0.00466927
2006-05-31 00:00:00	3071.05	-0.0143862	9526	0.000630054	3284	0.0196808

Fig. 6 Original Crude Oil dataset

Index	Crude_Close_Price
0	3115
1	3342
2	3282
3	3283
4	3231
5	3205
6	3220
7	3284
8	3269

Fig. 7 Extracted Crude Close Price

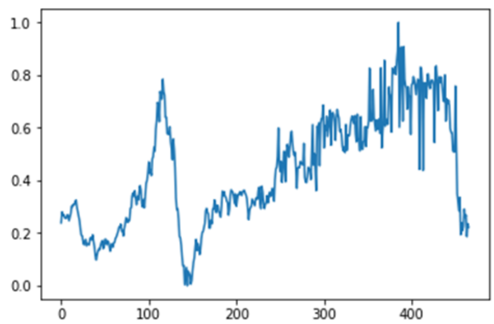

Fig. 8 Crude Oil dataset

Afterwards, three LSTM layers with lookback of 10 and 100 epochs were used and training score obtained was 269.17 RMSE and testing score obtained was 566.34 RMSE. These results were very close to the desired results. The obtained results are shown in Fig. 11. Then, four LSTM layers with lookback of 10 and epoch 50 were employed in the proposed network. The obtained training score was 283.34 RMSE and the testing score was 532.13 RMSE. The reduction in the number of epochs to 50 leads to the encouraging results. The obtained results are depicted in Fig. 12.

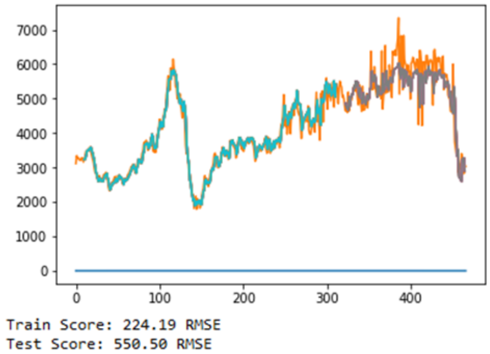

Fig. 9 Results with 10 lookback, 100 epochs and 4 LSTM layers

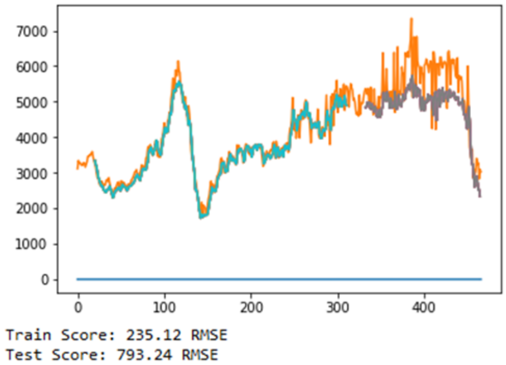

Fig. 10 Results with 20 lookback, 100 epochs and 6 LSTM layers

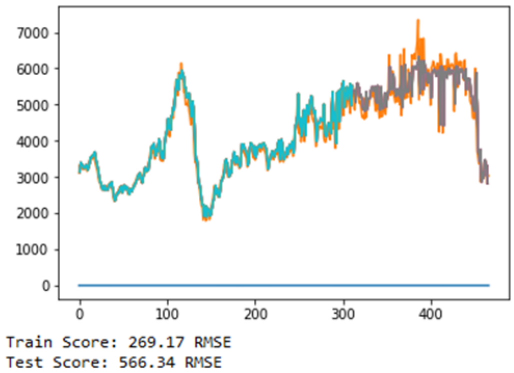

Fig. 11 Results with 10 lookback, 100 epochs and 3 LSTM layers

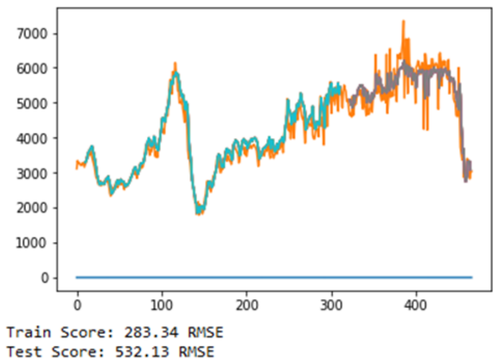

Fig. 12 Results with 10 lookback, 50epochs and 4 LSTM layers

The various results obtained from experimentation with the proposed network are shown in Table I. In the end, it was concluded that lookback of 10 and 4 LSTM layers provided the best results and thus were taken as the final configurations of the proposed architecture of the network. Table I summarizes various results obtained after experimentation with the proposed network.

TABLE I
RESULTS OBTAINED FOR VARIOUS CONFIGURATIONS OF THE PROPOSED NETWORK

LAYERS (LSTM)	LOOK BACK	EPOCH	TRAIN (RMSE)	TEST (RMSE)
4	1	100	240	608
4	20	200	226	727
4	20	100	247	824
4	10	100	224.19	550.5
6	20	100	235.12	793.2
3	10	100	269.1	566.3
4	10	100	283	532

V. CONCLUSION AND FUTURE WORK

Crude oil price forecasting plays a significant role in world economy and its accurate prediction has significant benefits for the economic conditions of a country. In this direction, an effort has been in this paper. This paper has proposed an LSTM based network for better prediction of crude oil prices. The results obtained from the work are quite encouraging. The results indicate that large lookups do not necessarily improve the accuracy of the predictions of crude oil prices. It has been found that lookups up to the value of 10 are ideal for crude oil price prediction purposes. It has also been found that just increasing the number of LSTM layers do not have much impact on the accuracy of the results. In future work, current market and political conditions can also be taken into consideration in crude oil price forecasting for even better results.

REFERENCES

[1] Mohammad Reza Mahdiani and Ehsan Khamehchi, “A modified neural network model for predicting the crude oil price”, Intellectual Economics, vol. 10, no. 2, pp. 71-77, Aug. 2016.
[2] Manel Hamdi and Chaker Aloui, "Forecasting Crude Oil Price Using Artificial Neural Networks: A Literature Survey," Economics Bulletin, AccessEcon, vol. 35, no. 2, pp. 1339-1359, 2015.
[3] Yu Runfang, Du Jiangze and Liu Xiaotao, “Improved Forecast Ability of Oil Market Volatility Based on combined Markov Switching and GARCH-class Model, Procedia Computer Science, vol. 122, pp. 415-422, 2017.
[4] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink and J. Schmidhuber, "LSTM: A Search Space Odyssey," IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 10, pp. 2222-2232, Oct. 2017.
[5] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
[6] Colah’s Blog: Understanding LSTM Networks. Accessed from http://colah.github.io/posts/2015-08-Understanding-LSTMs
[7] S. Moshiri, and F. Foroutan, “Forecasting nonlinear crude oil futures prices,” The Energy Journal vol. 27, pp. 81-95, 2005.
[8] Siddhi Vinayak Kulkarni and Imad Haidar, Forecasting Model for Crude Oil Price Using Artificial Neural Networks and Commodity Futures Prices. International Journal of Computer Science and Information Security, vol. 2, no.1, June 2009.
[9] Manel Hamdi and Chaker Aloui, “Forecasting Crude Oil Price Using

Artificial Neural Networks: A Literature Survey", Economics Bulletin, vol. 35, no. 2, pp. 1339-1359, 2015.

[10] S. N. Abdullah and X. Zeng, "Machine learning approach for crude oil price prediction with Artificial Neural Networks-Quantitative (ANN-Q) model," The 2010 International Joint Conference on Neural Networks (IJCNN), Barcelona, pp. 1-8, 2010.

[11] Yanhui Chen, Kaijian Heb, and Geoffrey K.F. Tso, "Forecasting Crude Oil Prices: a Deep Learning based Model", Procedia Computer Science, vol. 122, pp. 300-307, 2017.

[12] Ilya Sutskever, Oriol Vinyals, and Quok V Le. "Sequence to sequence learning with neural networks. Advances in neural information processing systems, pp. 3104–3112, 2014.

[13] Ji Young Lee and Franck Dernoncourt. Sequential short-text classification with recurrent and convolutional neural networks. arXiv preprint arXiv:1603.03827 (2016).

[14] Y. Cheng, C. Xu, D. Mashima, V.L.L. Thing, and Y. Wu, "Power LSTM: Power Demand Forecasting Using Long Short-Term Memory Neural Network.", In: Cong G., Peng WC., Zhang W., Li C., Sun A. (eds) Advanced Data Mining and Applications. ADMA 2017. Lecture Notes in Computer Science, vol 10604. Springer.

[15] K. Chen, Y. Zhou, and F. Dai, "An LSTM-based method for stock returns prediction: A case study of China stock market," in Big Data (Big Data), 2015 IEEE International Conference on, Oct 2015, pp. 2823–2824.

[16] Di Wang and Eric Nyberg," A Long Short-Term Memory Model for Answer Sentence Selection in Question Answering". In ACL (2), pp. 707–712, 2015.

[17] Yu Kang Jia, Zhicheng Wu, Yanyan Xu, Dengfeng Ke, and Kaile Su, "Long Short-Term Memory Projection Recurrent Neural Network Architectures for Piano's Continuous Note Recognition," Journal of Robotics, vol. 2017, Article ID 2061827, 7 pages, 2017.

[18] K.L. Chu and K.S.M. Sahari, "Behavior recognition for humanoid robots using long-short-term memory", International Journal of Advanced Robotic Systems, vol. 13, no.6, pp. 1-14, 2016.

[19] Minlie Huang, Yujie Cao, and Chao Dong, "Modeling rich contexts for sentiment classification with LSTM". arXiv preprint arXiv:1605.01478. 2016.

[20] "Crude Oil Dataset" accessed online from https://www.kaggle.com/victor7246/weeklydatacommodity on 25-Dec-2017.

[21] Moalla H., Elloumi W., Alimi A.M. (2017) H-PSO-LSTM: Hybrid LSTM Trained by PSO for Online Handwriter Identification. In: Liu D., Xie S., Li Y., Zhao D., El-Alfy ES. (eds) Neural Information Processing. ICONIP 2017. Lecture Notes in Computer Science, vol 10637. Springer, Cham.

Varun Gupta received his PhD in Computer Engineering from National Institute of Technology, Kurukshetra (Haryana), India. He is currently an Associate Professor in the Department of Computer Science and Engineering, Chandigarh College of Engineering and Technology (Degree Wing), Chandigarh, India. Before joining Chandigarh College of Engineering and Technology, he worked in the power sector as an Information Technology Engineer. He has published several technical research papers in leading journals and conferences from IEEE, Elsevier, Springer etc. His research interests include Artificial Intelligence, Deep Learning, Machine Learning, Big Data Analytics, Smart Grid and Software Engineering.

Ankit Pandey is currently pursuing Bachelor of Engineering in computer science and engineering from Chandigarh College of Engineering and Technology, Chandigarh, India. His main research interests include neural networks and machine learning.