

Crude Oil Prediction Using LSTM

Nidhi Moitra, Priya Raj, Sanidhya Saxena, Rohit Kumar

Department of Information Science and Engineering Dayanand Sagar Academy of Technology and Management Bengaluru, India

Abstract:- Crude oil has another name called black gold which has an essential role in evolution of global wealth and financial market. Therefore, dynamic information of future expected price will lead to enhancement of decision making at different levels. Specifically, this is an attempt made to forecast price prediction using long short-term memory neural network rather than using convolutional neural network. We have come across testing different versions of model using various lookback and alternative tuning methods. The conclusion derived from this study are promising and represent a more precise prediction for the crude oil price in coming days.

Keywords:- Crude Oil Prediction, LSTM, Recurrent Neural Networks.

I. INTRODUCTION

Crude oil has an essential role in the world, as this is one of the major products worldwide and thus includes global measurements. The origin of crude oil prediction errors involves composite supply-demand structures. Crude oil volatility has a critical effect on economic factors which includes economic increase of the country, unemployment, exchange rate that moreover depends on crude oil export and import. In current days machine learning techniques can be implemented in various applications. Machine learning gives powerful computational tools and algorithms that are capable of learning itself and predict data with long short-term memory. This paper contains LSTM based recurrent neural networks for the matter of crude oil price prediction. Recurrent neural networks (RNN) identifies to be the most powerful and impactful models for processing time-series based sequential data. LSTM variants can be used for other task as well other than prediction such as speech, handwriting and polyphonic modelling. The hyperparameters of variants were assessed using random search and analysed using variance framework. ANN consist of set of nodes which resembles to the neurons of the biological brain. The mission of ANN was to solve a task in a same way the human brain would do . ANN is rarely applied in prediction of model as sometimes tries to over-fit the relationship. It is also used in such cases where the past will likely to appear in future in the same way.

Backpropagation is used for improving the precision of prediction in field of data mining and machine learning. It computes the slope of the loss function based on the neural network parameters Another approach to oil price prediction is the use of decision tree approach. It helps us not only with prediction and classification, but also a very powerful tool to know the behaviour of different variables. Decision Tree is a part of superintend learning algorithm that works for categorial input and output variables. The input values or attributes can be the economic indicators that effect the prediction of crude oil prices.

II. SURVEY STUDY

Many Scientist and researcher have come across unique and variation of model for discovering and exploring and forecasting crude oil prices . Apprehensive study related to the forecasting and prediction of prices to be contingent on economic models. Also more study about intelligent and smart algorithm is taken into account. In the short-term it was Observed , variance was inflating by the events which were irregular. One of the feature of the imperfection of all the methodology was that the upcoming movement of oil price were derived from the prior data. A lot of machine learning strategy came into existence for oil price prediction which can be defined by ANN. Scientist have perceived that discussion of a topic done online convey point of view or outlook based on a subject matter , and thus topic models may have consequences on market prices. But in most cases, machine learning techniques compared to the conventional techniques varies on permanent set of training data to coach a machine learning model followed by series of solicitation of model to the test set. One of the prime data source defines the important and useful information to be inserted in unstructured big data such as data n form of text which can employed for oil cast forecasting. In modern times , a number of studies have made essential offering for prediction in retail in the sphere of text mining. Convolutional neural network is a neural network based on deep learning concept. Its application has been mostly in recognition of image and speech and sentence modelling. The validation from the EMD technique propose regarding the prime cause behind the fluctuation of price is midterm weighty incident. When out of the blue and unplanned events crash a market , it results from decomposition integration method which decline performance.

III. METHODOLOGY

A. System design

The main aim of this structure incorporated in study can fetch out data from economical news and propose this sets into prognosticate model. Major phases in formulated system include data collection and pre-processing, feature and factor selection and price appraisal and prediction. In the initial hand, news, financial and market data are gathered and processed. In Further aspect, unstructured documents are modified into structured extract by CNN classification.

B. Data retrieval and pre-processing

In data retrieval, datasets can be fetched such of news data, black gold price data and market data. Dataset from news can be retrieved through headlines as it is easier to obtain and justifies in one line. Factors that affect the prediction are exportbusiness, stock market and later business.

C. Sentimental Analysis

In this era of modernization, big data is also assisting through study of sentiment analysis which focuses on retrieving data through news and proposing prediction model. In this kind of analysis dictionary-based approach is accounted to gather the data regarding markets and essential factors affecting it. In case of trend prediction, the sentiment and prediction models are considered as variables.

D. Back Propagation

Back-propagation is considered as an algorithm which can be used for the purpose of training feed forward neural networks for prognosticate learning model. This leads to the attainably use gradient methods to teach multi – layer networks, by modifying weights to minimum loss. The process fetch the inputs and outputs and modify its inner state that will be capable enough to calculate the output that will be very precise to the expected output.

Backpropagation can also be described as "backward propagation of errors." It is a natural function to teach artificial neural networks.

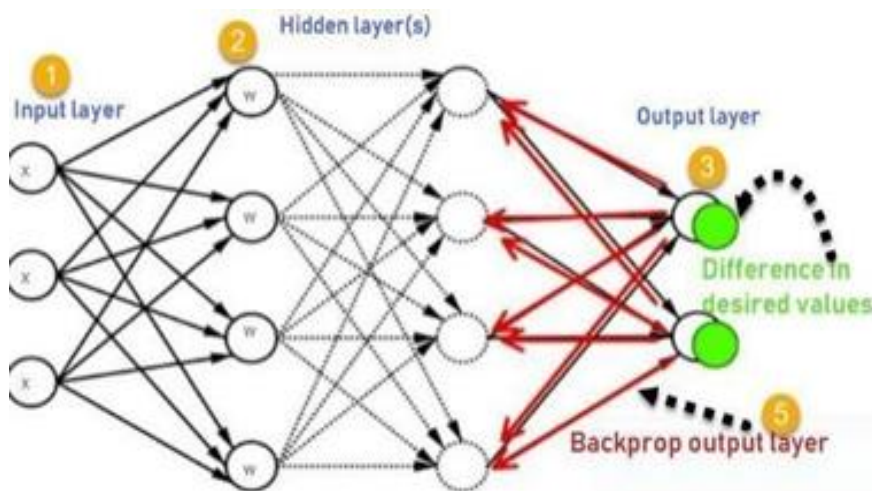


Fig 1:- Back Propagation in neural networks.

E. CNN (convolution neural network)

Also known as ConvNet is a type of neural network using convolution and pooling layer can be used for time series prediction. Convolution layer is responsible for passing the result to next layer. Pooling layer in the other hand can be used to decrease the dimensionality or drop the amount of parameters. The forecasting can be done with TensorFlow which is followed by getting the data,

generating features , generating ML model , training the ML model and testing and predicting with ml model. Movement of the price which can fall and rise can be termed as the outcome of CNN classification. Generally the activity of price can be specified by:

$$M_t = \{0, p_t < p_{t-1}\} \{1, p_t \geq p_{t-1}\}$$

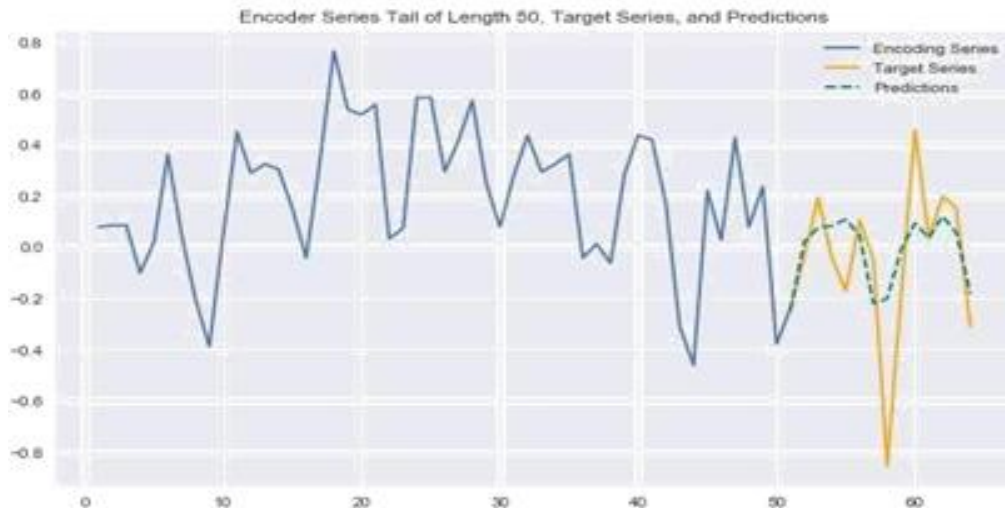


Fig 2: Time Series Forecasting with CNN

F. RNN (Recurrent Neural Network)

RNN are not same as feedforward networks. To predict things, it uses internal memory. It is capable of doing things which humans can't do such as handwriting, speech recognition. Sequential information are added as

input to these networks.

We assume inputs are independent of each other which is a false assumption. We should know the previous words so that next words can be predicted without any trouble.

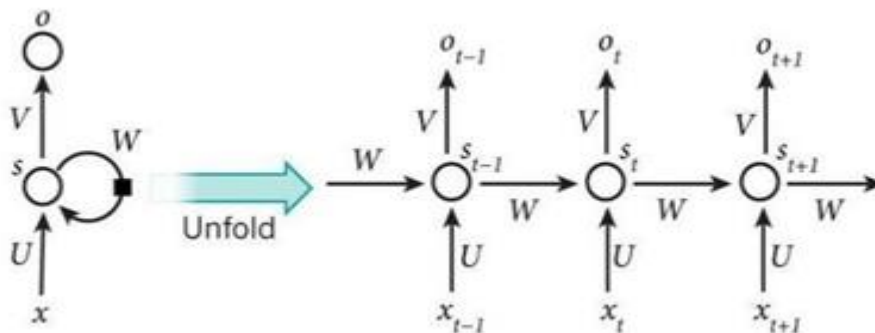


Fig 3:- RNN

One extension of RNN is bidirectional RNN. In this at time t the output also depends on future inputs.

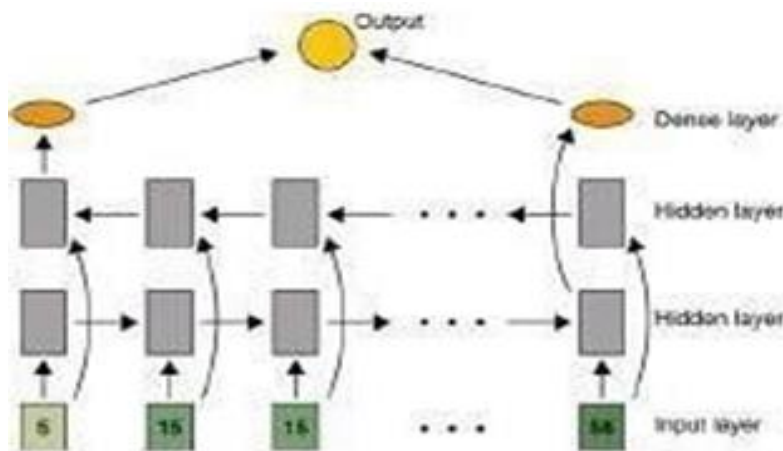


Fig 4:- Deep RNN

Another extension is deep RNN. It consists of multiple layers per time step.

IV. RESULT DISCUSSION

To decide the concluding architecture of the given network, different designs were tested. The following table represents the test-cases.

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

Table 1:- Represent some results got from procedure developed with define network is shown above. At last,it finishes the suitable results and hence taken as the final configuration of the suggested architecture of the network.

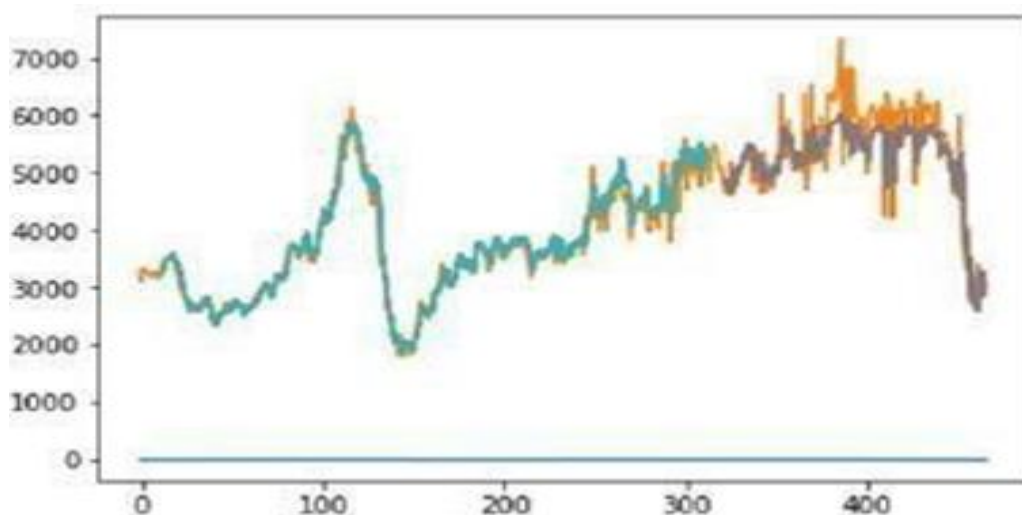


Fig 5:- Results with 10 look back,100epochs and 4LSTM layers.

Train Score:224.19 RMSE Test Score:550.50 RMSE

Therefore figure 5 explains the graph representation of the data we took that includes 10 lookback,100 epochs and 4 lstm layers.

As a result we obtained train score :224.19 RMSE and Test Score :550.50 RMSE.

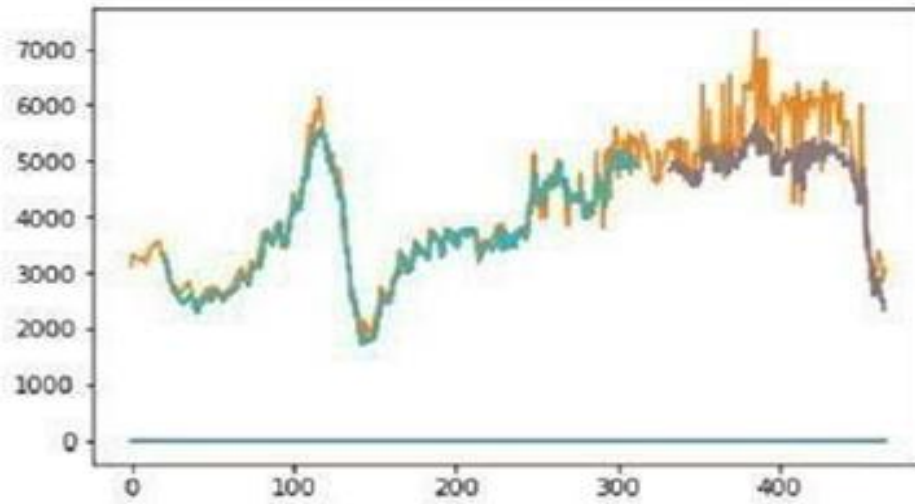


Fig 6:- Results with 20 look back,100epochs and 6LSTM layer.

Train Score:235.12 RMSE Test Score:793.24RMSE

Here in figure 6 the graph explains the data which conatins 20 lookback , 100 epochs and 6 lstm layers. As a result we can conclude that the values are as follows:

The Train score is: 235.12 RMSE The Test score is 793.24 RMSE

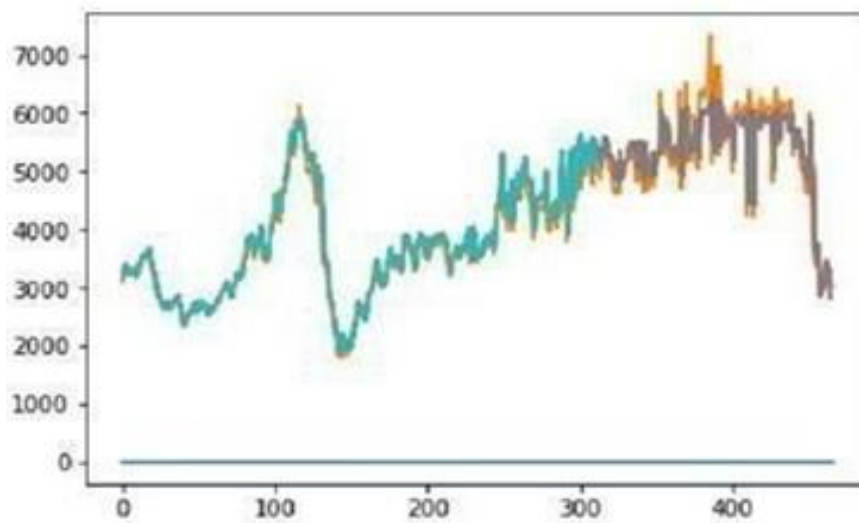


Fig 7:- Results with 10 look back,100epochs and 3LSTM layer.

Train Score:269.17 RMSE Test Score:566.34 RMSE

The figure 7 explains the data which includes 10 look back , 100epochs and 3 lstm layers which shows the train score to be 269.17 RMSE and Test score to be 566.34 RMSE. Fig 10: Results with 10 look back,50 epochs and 4 LSTM layer.

It is clear from the above results that 4LSTM with 10 lookbacks and 50 epochs is best suited for this particular problem statement. Hence proved that LSTM is a better model compared to traditional RNN feed forward model for recurring neural networks to predict the crude oil prices.

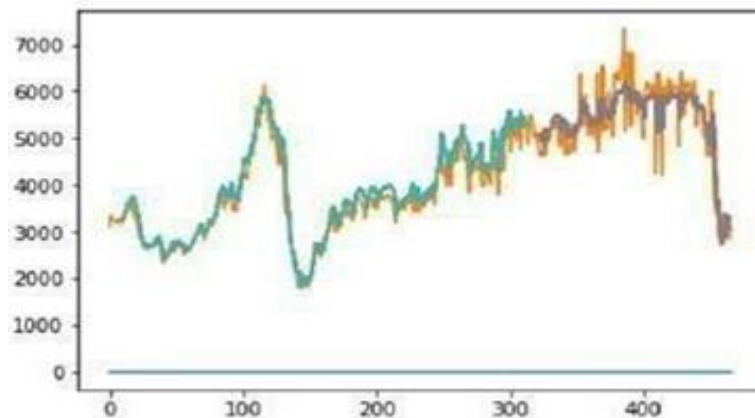


Fig 8:- Result with 10 lookback, 50 epochs and 4lstm layer.

The last test-case explains the graph with data including 10 lookback, 50 epochs and 4 lstm layers which integrate the result and the train score is: 283.34 RMSE and Test score is : 532.13 RMSE.

V. CONCLUSION

This paper has cleared that an LSTM network is better than other traditional neural network for forecasting prices as it aims in using backpropagation model. Traditional neural network such as rnn or cnn on the other hand predicts the next outgoing but doesn't necessarily save the previous data or connection which is based on feed-forwarding, in the sense the previous data is not necessary to predict the future data. LSTM focuses on storing the previous data and prediction which is rather encouraging and more approximate. The outcome derived are relatively encouraging. The results show that large look ups do not necessarily improve the accuracy of the predictions of crude oil prices. Hence it can be concluded, the model with single LSTM model is definitely the most accurate.

REFERENCES

- [1]. Varun Gupta, Ankit Gupta ;Crude Oil Price Prediction Using LSTM Networks World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering Vol:12, No:3, 2018.
- [2]. Text Based Crude Oil Price Forecasting; A Deep Learning Approach Xuerong Li(a) Wei Shang (b)(a) Shouyang Wang(a)(b) (a)School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China (b) Academy of Mathematics and System Science, Chinese Academy of Sciences, Beijing 100190, China.IEEE(2018)
- [3]. A Decision Tree Approach To Oil Price Prediction ;Nnamdi I. Nwulu Department of Electrical & Electronic Engineering Science University of Johannesburg
- [4]. Long Short Memory Networks For Automatic Generation Of Conversation; Tomohiro Fujita Wenjun Bai Changqin Quan Department of Computational Science Graduate School of System Informatics, Kobe University Kobe, Japan @cs11.cs.kobe-u.ac.jp.
- [5]. quanchqin@gold.kobe-u.ac.jp.IEEE(2017)
- [6]. A Critical Review On The Use Of Artificial Neural Networks in Olive Oil Production characterization and authentication; I. Gonzalez-Fernandez, M. A. Iglesias-Otero, M. Esteki, O. A. Moldes, J. C. Mejuto & J. Simal- Gandara.IEEE(2018).
- [7]. Artificial Neural Network Based on Oil Price Forecasting. A decade Review Of the Literature; Mandeep Kaur¹, Parminder Kaur² ^{1, 2} Department of Computer Science, Khalsa College for Women, Civil Lines, Ludhiana. Volume 4 Issue IX, IEEE September 2016.
- [8]. Forecasting Weekly Crude Oil Using Twitter Sentiment of US Foreign Policy and Oil Companies Data; A. Zaidi Natural Language Processing Group, University of Cambridge, Cambridge CB3 0FD, UK Ahmed.Zaidi@cl.cam.ac.uk M. Oussalah Centre for Ubiquitous Computing, Faculty of Information technology, University of Oulu, Finland. Mourad.Oussalah@oulu.fi.IEEE(2018).
- [9]. Ensemble Incremental Random Vector Functional Link Network For Short-Term Crude Oil Price Forecasting; Xueheng Qiu and P. N. Suganthan School of Electrical and Electronic Engineering Nanyang Technological University, Singapore {xiux0004, epnsugan}@ntu.edu.sg.IEEE(2018).
- [10]. Extracting Knowledge From Technical Reports For The Valuation of West Texas Intermediate Crude Oil Futures; Joseph D. Prusa*, Ryan Sagul, Taghi M. Khoshgoftar*, Michael Sterling Florida Atlantic University {jprusa, khoshgof}@fau.com {ryansagul, msterli6}@gmail.com.IEEE(2017).
- [11]. Mohammad Reza Mahdiani and Ehsan Khamnehchi, "A modified neural network model for predicting the crude oil price", Intellectual Economics, vol. 10, no. 2, pp. 71-77, Aug. 2016.
- [12]. 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.