Petroleum 6 (2020) 353–361

Contents lists available at ScienceDirect

Petroleum

journal homepage: http://www.keaipublishing.com/en/journals/petroleum

CPPCNDL: Crude oil price prediction using complex network and deep learning algorithms

Makumbonori Bristone[a], Rajesh Prasad[b,*], Adamu Ali Abubakar[a]

[a] School of IT and Computing, American University of Nigeria, Yola, Adamawa State, Nigeria
[b] Department of Computer Science and Engineering, African University of Science and Technology, Abuja, Nigeria

ARTICLE INFO

Keywords:
Complex network analysis
Deep learning
Long-short term memory network
K-core centrality
Artificial intelligence
Crude oil price prediction

ABSTRACT

Crude oil price prediction is a challenging task in oil producing countries. Its price is among the most complex and tough to model because fluctuations of price of crude oil are highly irregular, nonlinear and varies dynamically with high uncertainty. This paper proposed a hybrid model for crude oil price prediction that uses the complex network analysis and long short-term memory (LSTM) of the deep learning algorithms. The complex network analysis tool called the visibility graph is used to map the dataset on a network and K-core centrality was employed to extract the non-linearity features of crude oil and reconstruct the dataset. The complex network analysis is carried out in order to preprocess the original data to extract the non-linearity features and to reconstruct the data. Thereafter, LSTM was employed to model the reconstructed data. To verify the result, we compared the empirical results with other research in the literature. The experiments show that the proposed model has higher accuracy, and is more robust and reliable.

1. Introduction

Crude oil is a major source of Energy in Nigeria and the entire world. Crude oil being the main backbone of the Nigerian economy plays a critical role in redefining the global economy as well as the Nigerian economy in particular. The impact of crude oil on Nigerian economy has been double edge [1]. Crude oil price fluctuations have consequential effect on the country's economic activities, social stability as well as national security [2], therefore proactive knowledge of its future fluctuations can lead to better decision making in numerous managerial levels. The Nigerian government's annual budget has always been pegged to a specific amount of the International price of crude oil, thus, making both the government fiscal and monetary policy to be susceptible to fluctuation that may arise in crude oil price volatility which engenders the performance of the economy through the exchange rate [3]. The age-long debate among economists and researchers regarding the effect of crude oil price fluctuations and economic performance still remains unresolved. Hence the need to predict price of crude oil accurately as much as possible is a major concern of government agencies, institutions and research communities. Crude oil price is among the most complex and tough to model because fluctuations of price of crude oil are highly irregular, nonlinear and varies dynamically with high uncertainty [4]. The odd behavior of crude oil price instability has been adjudged due to some factors such as demand and supply in the financial market, economic growth, technological development that build complex relationship between these factors and the oil price [5]. Predicting prices of crude oil accurately is one of the most fascinating and gripping issues faced by economist and academic researchers these days, due to the fact that crude oil is one of the major sources of energy worldwide and its price movement can disturb aggregate economic activities. Prior knowledge of prices of crude oil can be considered to be among the key parameters needed to make a proper decision towards development, production processes and government short term and long-term planning, hence any rise or fall in crude oil price has a measurable effect on the economy [6].

Due to the immeasurable impact of crude oil to the Nigerian economy and the world at large, the need on how to proactively predict accurately the crude oil prices for effective planning, budgetary purpose and other government essential activities cannot be over emphasized. Studies by Ref. [7] has demonstrated a relationship between oil price and gross domestic product (GDP) growth rate and asserted that the significant impact on the economy can be observed only through a high increase in the price of oil. However, prices of crude oil displaying complex nonlinear characteristics and variability are very hard to predict because of the intrinsic difficulty associated with the data set. As such, seeking for a promising predictive approach for crude oil price

Peer review under responsibility of Southwest Petroleum University.
* Corresponding author.
E-mail address: rprasad@aust.edu.ng (R. Prasad).

https://doi.org/10.1016/j.petlm.2019.11.009
Received 19 June 2019; Received in revised form 24 November 2019; Accepted 25 November 2019
Available online 27 November 2019

series is hardly outdated since crude oil is the main source of energy in the world and predicting oil price still remains a major bottleneck.

Determining a model for predicting several economic parameters is a very common approach and this is not an exception to crude oil price prediction [8]. There are several research works that have attempted to predict the prices of crude oil some of which include: Researchers in Ref. [9] examine crude oil prices by applying a three-layer feed forward neural network and genetic algorithms (GA). The result shows that, the optimized model proved to be more powerful against traditional artificial neural network (ANN) as regards to accuracy, variability, fitness and precision. Researchers in Ref. [10] examine movement of price of crude oil by using modified semi supervised learning (SSL) algorithms. This approach modified the prevailing SSL algorithm for utilization of time series algorithms. The technique not only influences the input variable over desired variable but also interactive influence among variables. The complexity as well as the irregularity of crude oil price was resolved by representing the oil prices as nodes in a network and the prediction was done by propagating effect of neighboring economic factors through connections, the result shows that this technique can predict more accurately when compared to prediction techniques offered by prevailing prediction models. Researchers in Ref. [11] examine the complexity of crude oil price data with regard to its interactive involving factors by applying a powerful strategy of "decomposition and ensemble" with a swift and steady artificial intelligence (AI) predictive approach (i.e. extended extreme learning machine (EELM)) to modeled the volatility and irregularity of the crude oil price. The findings reveal that, the ensemble learning model remarkably improved prediction performance and statistically performed better than other forecasting techniques in terms of accuracy, time saving and robustness. Researchers in Ref. [5] examine the complex correlation between crude oil price with several other factors by applying a deep learning assemble method called SDAE-B to forecast crude oil price. In their research, the SDAE-B out performs benchmark models including traditional econometric models (i.e. random walk (RW) and Markov regime switching (MRS)), machine learning models of shallow architectures (i.e. feedforward neural network (FNN), support vector regression (SVR) and statistic test confirm the supremacy of the model. Researchers in Ref. [12] applied the deep learning technique in forecasting crude oil price, they particularly used two specific deep learning models, deep belief network (DBN) and recurrent neural network in modelling the nonlinear characteristics in crude oil price volatility.

The superiority performance of DBN based model against the benchmark autoregressive moving average (ARMA) model is striking; however, it is not statistically impressive when the predictive power is compared against the RW model. This leads to the discovery that the predictive performance of deep learning is very sensitive to parameters.

The most obvious difficulty faced by numerous researchers in modelling crude oil prices is the highly complex nature of the dataset resulting to excessive noise in the real data and decreases the strength of the prediction accuracy of the models [13]. Research conducted by Refs. [14,15] portrays that, noise reduction techniques includes the following methods, wavelength decomposition (WT), singular spectral analysis (SSA), empirical mode decomposition (EMD), variational mode decomposition (VMD), entropy-based wavelet de-noising, hybrid stanlet de-noising based on the least squares support vector regression model, exponential smoothing based on neural networks, and the extended Kalman filter method. However, all these techniques have their own flaws parameter sensitivity and time consumption are their major drawbacks which can be overcome by using the idea of complex network analysis.

Thus, in this paper, a deep learning approach known as long short-term memory (LSTM) network was proposed to predict the price of crude oil. LSTM was established to mitigate the important issue with standard recurrent neural network's (RNNs) difficulty in learning long range dependencies between data instance that are far from each other. Basically, this approach is an evolution of a standard RNN where hidden layer is replaced by a *memory cell* [16]. A memory cell contains a node with a self-connected recurrent edge, ensuring that the gradient can pass across many time steps without vanishing. The concept of LSTM is to ensure that long-range dependencies by introducing an intermediate storage within the memory cell which is being controlled by special neurons call gates. Regardless of their simplicity, LSTMs approach surprisingly performs excellently in numerous tasks and either they or their variants represent the most widely adopted RNNs application [17–20]."

Thus, considering the important role crude oil price plays in the economy today, the greater demand of techniques that will efficiently and proactively predict accurately future behaviors of crude oil price justifies the need for effective, prediction approaches, hence the reason for adopting LSTM in this paper.

In general, the paper makes two major contributions, firstly, a novel visibility graph algorithms and k-core centrality was applied to preprocess and reconstruct the data. Secondly, the LSTM was developed to model and predict future behavior of the crude oil price and then ten (10) crude oil prices were selected across the world to evaluate the effectiveness and performance of the model.

The remaining part of this paper is organized as follows: Section 2 describes related concepts on crude oil price prediction. Section 3 describes data preprocessing process, formulation process of the model based on the LSTM algorithms of deep learning. The result and analysis of model, evaluation of the effectiveness of the model based on 10 different crude oil prices across the world were discussed in Section 4. Section 5 discussed conclusions and future research.

2. Related concepts

Crude oil price prediction is a wide area of research that has been on for a very long time in history and numerous approaches have been proposed in predicting crude oil price. Research works on crude oil price prediction approaches can be broadly categorized in three: *statistical and econometric models, AI* and *hybrid modeling techniques* [5,11,21].

The statistical and econometric models cover many familiar models: Research in Ref. [2] analyze the composition of crude oil price based on empirical mode decomposition (EMD) by replacing EMD with extended empirical mode decomposition (EEMD). Authors in Ref. [22] uses a dynamic model averaging approach (DMA) to model and forecast crude oil price on west Texas intermediate (WTI) crude oil market, the results show that, DMA technique provides a better proxy of expected spot price than future price. Authors in Ref. [23] applied independent component analysis (ICA)-based support vector regression scheme (ICA-SVR) to identify the underlying factors responsible for volatilities in price of crude oil. Authors in Ref. [24] applied different versions of generalized autoregressive conditional heteroscedastic (GARCH) models such as GARCH, extended generalized autoregressive conditional heteroscedastic (EGARCH), integrated generalized autoregressive conditional heteroscedastic (IGARCH), fractionally integrated generalized autoregressive conditional heteroscedastic (FIGARCH), glosten-jagannathan-runkle generalized autoregressive conditional heteroscedastic (GJR-GARCH), and fractionally integrated asymmetric power autoregressive conditional heteroscedastic (FIAPARCH), to determine which one of these models performs more accurately in terms of ingroups and intergroup activities. Authors in Ref. [25] use vector autoregressive model (VAR) and vector error correction model (VECM) to explore relationships between price of crude oil, corn and ethanol. The result shows that the development of ethanol production is one way of controlling inflation in relation to crude oil prices. Authors in Ref. [26] employ a structural VAR model of the global crude oil market and United States (US) corn market. The result illustrates that close to 36% of the variation in the real price of corn can be attributed to structural supply and demand shocks in the global crude oil market. Table 1 summarizes prediction literature by means of statistical and

Table 1
Literature using statistical and econometric models for predicting price of crude oil.

Typical Literature	Prediction Models	Main Result
[2]	EEMD	Replaces EMD with EEMD
[22]	DMA Model	The proposed DMA Model provides a better proxy of expected spot price than future price.
[23]	ICA-SVR	Their prediction models identify factors underlying volatilities in crude oil prices.
[24]	Different versions of GARCH models	For two periods, asymmetric and integrated GARCH models provide relatively more accurate performance than the other available models.
[25]	VAR and VECM	The empirical result shows that the development of ethanol production is one way of controlling inflation in relation to crude oil prices.
[26]	VAR	The empirical result shows that close to 36% of the variation in the real price of corn can be attributed to structural supply and demand shocks in the global crude oil market.

econometric models.

In essence, because the econometric models assume the data to be stationary, regular and linear, they cannot accurately model time series data that are irregular, complex and nonlinear.

In addition to the traditional statistical and econometric models, AI techniques have been used to uncover the inner complexity of prices of crude oil. For instance Ref. [27], applied support vector machine (SVM) which has ability of modelling the nonlinearity and to prove that it could be used in predicting crude oil price with high degree of accuracy. Researchers in Ref. [28] proposed a dynamic correcting support vector regression machine in predicting prices of crude oil, the result obtained was good, it could be easily used in predicting price of crude oil in our life. Researchers in Ref. [29] applied Firefly algorithm and least squares support vector regression (FA-LSSVR) to forecast crude oil price, the result obtained shows that the model outperforms single benchmark in terms of accuracy and time saving and robustness, suggesting that it is a promising alternative in forecasting price of crude oil. Researchers in Ref. [7] applied gene expression programming (GEP) and ANN through the interconnected neurons and training nodes learned from the given data and captures the nonlinear functional relations which were used as models to forecast oil price. Researchers in Ref. [10] predicted movement direction (upward and downward movement) in crude oil price using semi supervised learning (SSL)."The approach proposed by Ref. [5] uses stack denouncing auto encoders (SDAE) and bootstrapping aggregation (bagging), an algorithm based on deep learning and ensemble learning based forecasting method to forecast crude oil price. Researchers in Ref. [30] formulated a crude oil price prediction technique using ideas and tools from stream learning, a machine leaning paradigm for inferring and analyzing continuous flow of non-static data. Table 2 shows the literature using artificial intelligence approach to predict crude oil price.

In the last few decades, research works focused on hybridization of machine learning models. It tells us that a hybrid model gathers all the models' advantages and offset all the drawbacks, which leads to wide usage of hybrid models. The challenge is that the traditional approaches assume the behavior of the dataset to be stationary regular and linear while AI as single models suffered from local minima and overfitting issues. Hybrid models proved to have better forecasting accuracies than their corresponding single machine learning techniques.

For example, researchers in Ref. [31] forecast oil price with an ANN based dynamic nonlinear autoregressive with exogenous input (NARX) model which is a multivariate forecasting model. Researchers in Ref. [32] propose a hybrid forecasting model in which the training data is first preprocessed by compressed sensing denoising and then it is used for training certain machine learning techniques including ANN and support vector regression (SVR). Authors in Ref. [6] proposes a hybrid crude oil price forecasting model in which the meta parameters of ANN are selected by GA. Researchers in Ref. [33] proposes an ensemble oil price forecasting model based on SVR, instance-based learning (IBL) and K-star, and the prediction is generated by taking an average of all the individual forecasts of these machine learning techniques. There are also ensemble models that first decomposed oil price series into several components and then combine the forecasts of each components generated by Neural Networks (NNs) [11,34,35]. Table 3 shows the literature using hybrid approach to predict crude oil price.

Empirical analysis results repeatedly demonstrate that hybrid prediction approaches prove to be more accurate than single techniques (ref. Table 3). This is the case because hybrid methods combine single models such that the advantages of one of the models compliment the disadvantage of the other(s). At the same time, the calculation process required in hybrid techniques is complicated. In other words, the hybrid prediction models are more likely to be advocated in recent literature, which also gives some hints for our research in this paper.

However, because of the high level of noise, complex nature and nonlinear characteristics within the crude oil dataset, this paper focuses on the following aim and objectives:

(1) To formulate a novel intelligent hybrid predictive model using complex network analysis and deep learning (DL).
(2) To enhance the performance of crude oil price prediction in Nigeria, as regards prediction accuracy, time saving and robustness, as well as do a comparative analysis on the performance of the model with other prediction models.
(3) And the specific objectives are:
    1) To preprocess data to remove noise and nonlinearity using visibility graph algorithm and k-core centrality of the complex network analysis.
    2) To develop a hybrid predictive model using the LSTM of the DL algorithm.
    3) To implement the proposed model on python programming environment and evaluate the performances of the models.

In order to accomplish the above set of objectives, we combined visibility graph algorithms and the k-core centrality of the complex

Table 2
Literature using artificial intelligence approach to predict crude oil price.

Typical Literature	Prediction Models	Main Result
[10]	SSL	SSL performs better than the ANN and SVM
[28]	Dynamic Correcting Support Vector Regression Machine	The experiment shows that the result is very good and could be easily used in predicting prices of crude oil in our life.
[7]	GEP	GEP Model performs better than the ANN and ARIMA
[5]	SDAE	SDAE performs better compared to traditional machine learning approach.
[30]	Stream Learning	Analysis and Inference for the continuous flow of non-stationary data

Table 3
Literature using hybrid approach to predict crude oil price.

Typical Literature	Prediction Models	Main Result
[34]	HTW-MBPNN	HTW-MBPNN performs better than the BPNN
[35]	EMD-SBM-FNN	EMD-SBM-FNN using the MIMO strategy is a very promising prediction technique with high quality forecasts and accredited computational loads for multi-step-ahead crude oil price forecasting
[11]	"EEMD-EELM"	EEMD-EELM is significantly superior to single EELM
[32]	"CSD-AI"	CSD-AI models outperform their single benchmarks in both level and directional predictions.
[6]	"GA-NN"	The GA-NN approach is able to improve prediction accuracy, and to simplify the complexity of the NN model structure.
[2]	A hybrid AI system framework	The proposed approach is significantly effective and practically feasible.

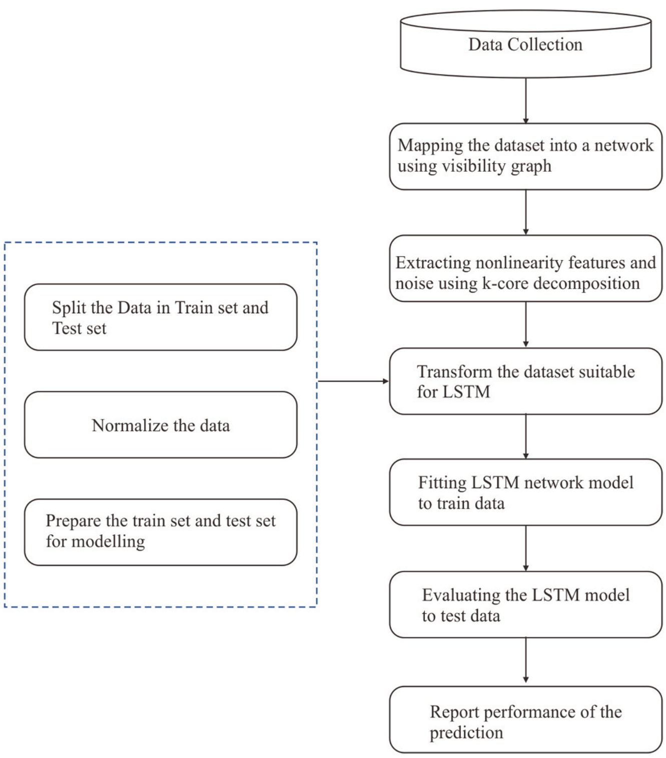

Fig. 1. Proposed methodology (prediction technique).

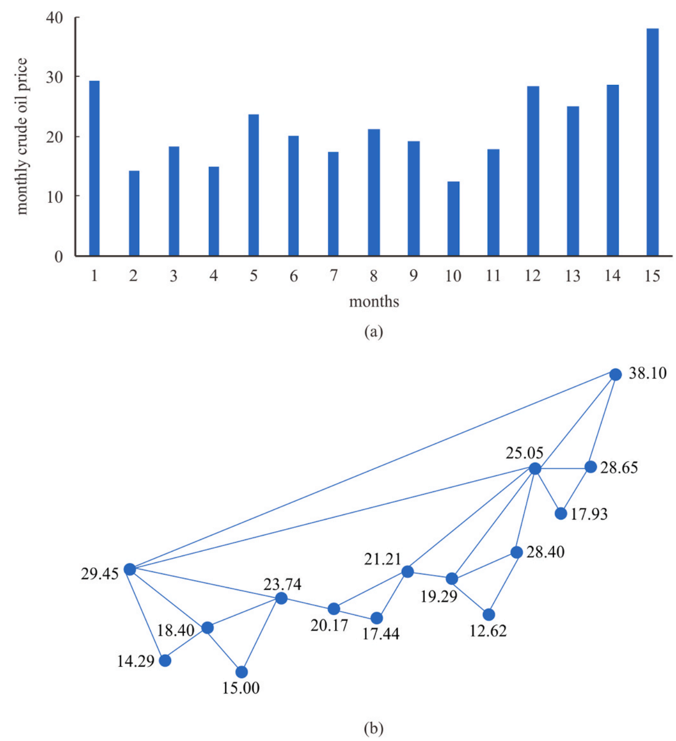

Fig. 2. (a)–(b). Mapped dataset of a visibility graph and Extracted network obtained.

network analysis with the LSTM of the deep learning in formulating the novel hybrid prediction model. To the best of our knowledge and belief, no studies has been carried out by combining the LSTM of DL with visibility graph and k-core algorithms of the complex network analysis in analyzing volatilities in crude oil price.

3. Materials and methods

The proposed prediction technique comprises of the following steps: mapping the datasets on a network using visibility graph algorithm, extraction of noise from the dataset and determination of the most influential nodes using k-core centrality, finally, LSTM is applied on the extracted datasets to train and test the models. At the end, the prediction of crude oil prices is evaluated with a view to discovering knowledge.

3.1. Prediction technique

The technique in Fig. 1 illustrates all the steps of the proposed technique and the implemented steps are explained in the sub-sections that follow.

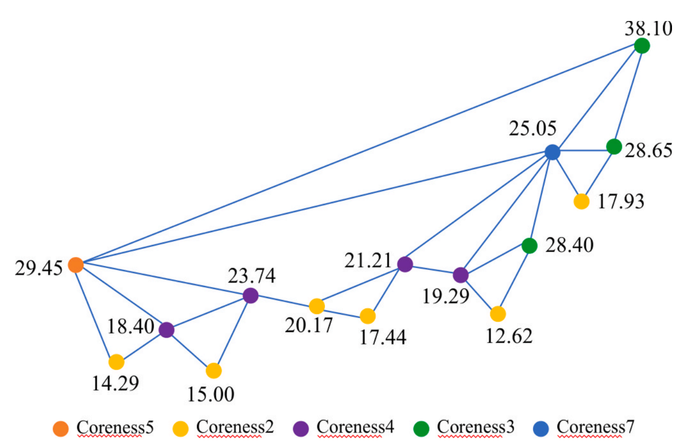

Fig. 3. k-core decomposition using k-core centrality algorithms.

3.2. Data preprocessing

Due to complexity, irregularity and non-linear nature of the data, predicting complex data such as crude oil price is one of the various tasks in machine learning. Complex network analysis has been widely

Table 4
Conversion of array of values into data matrix.

X	Y
0	14.29
14.29	15.00
15.00	20.17
20.17	17.44
17.44	12.62

Table 5
Transformation of time series to stationary.

Month	Y
1	14.29
2	15.00
3	20.17
4	17.44
5	12.62
Month	Y
0	0.71
1	5.17
2	−2.73
3	−4.82
4	2.2

Table 6
Transformed data to scale of 0–1.

Month	Scaled
0	0.22
1	0.31
2	1
3	0.64
4	0

Table 7
Crude oil price datasets summary.

Body	No. of years	No. of months
OPEC	1983–2013	360
CBN	2014–2018	56

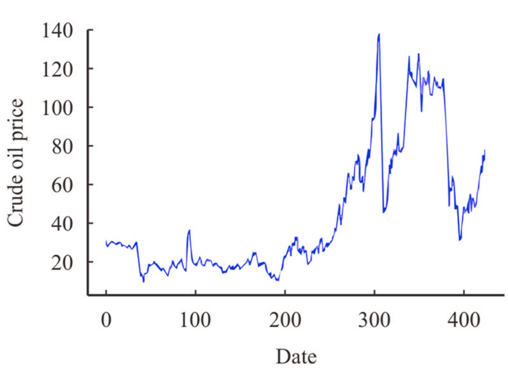

Fig. 4. Fluctuations of the crude oil price.

utilized recently to analyzed time series data and has proven to be very effective in yielding high quality result [36,37]. The approach involves two steps, the first step employs visibility graph algorithm to map the dataset into a complex network. The second step involves extracting the non-linearity features by applying k-core centrality."

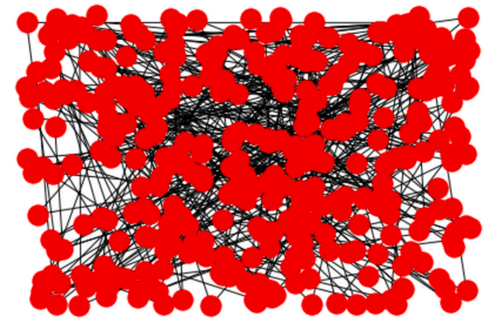

Fig. 5. Network of the dataset mapped on a network.

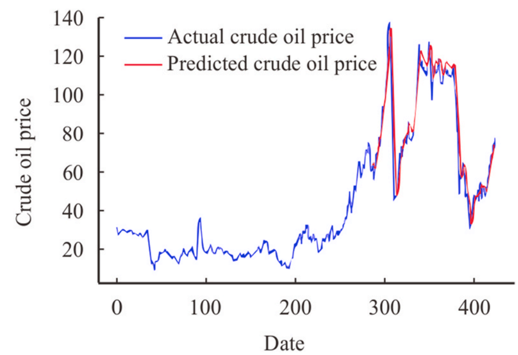

Fig. 6. Predicted and the actual crude oil price.

The algorithm to be used in mapping the datasets in a network is called visibility graph [38–40]. The idea is illustrated as follows: Assuming we have a time series data $X(t) = [x_1, x_2 \ldots x_n]$ of length n. Individual data point x_n in the series can be considered as vertex in the linked network and an edge can be drawn connecting to two vertices such that the two points are mutually visible with each other in the vertical bar time series. On the other hand, two data points are connected if a straight 'visibility line' joins the two points without crossing each other's intermediate data bar.

More formally the criteria can be established as: two data values x_a (at time t_a) and x_b (at time t_b) are connected if there exist a value (x_c, t_c) existing between the two (i.e.: $t_a < t_c < t_b$), that satisfied the following condition:" $x_c < x_a + (x_b - x_a)\frac{tc - ta}{tb - ta}$

The associated graph extracted is always:

(1) Connected: each node can at least be visible to its nearest neighbor (right or left)"
(2) Undirected: based on the buildup of the algorithm, the direction of the link is not defined"
(3) The visibility algorithm is invariant under rescaling of both horizontal and vertical axis."

For instance, given a price of crude oil as 29.45, 14.29, 18.40, 15.00, 23.74, 20.17, 17.44, 21.21, 19.29, 12.62, 28.40, 17.93, 25.05, 28.65, 38.10 in US Dollar per barrel. The result will indicate as contained in Fig. 2a and Fig. 2b.

3.3. Extracting the nonlinearity features using the k-core centrality

k-core centrality considers the number of nearest neighbor and asserts that the nodes with the same degree have the same influence in the network. Researchers in Ref. [41] argue that the position of node is more influential than its immediate neighbors. That is to say, if a node is located at the core part of the network, the influence of the node will

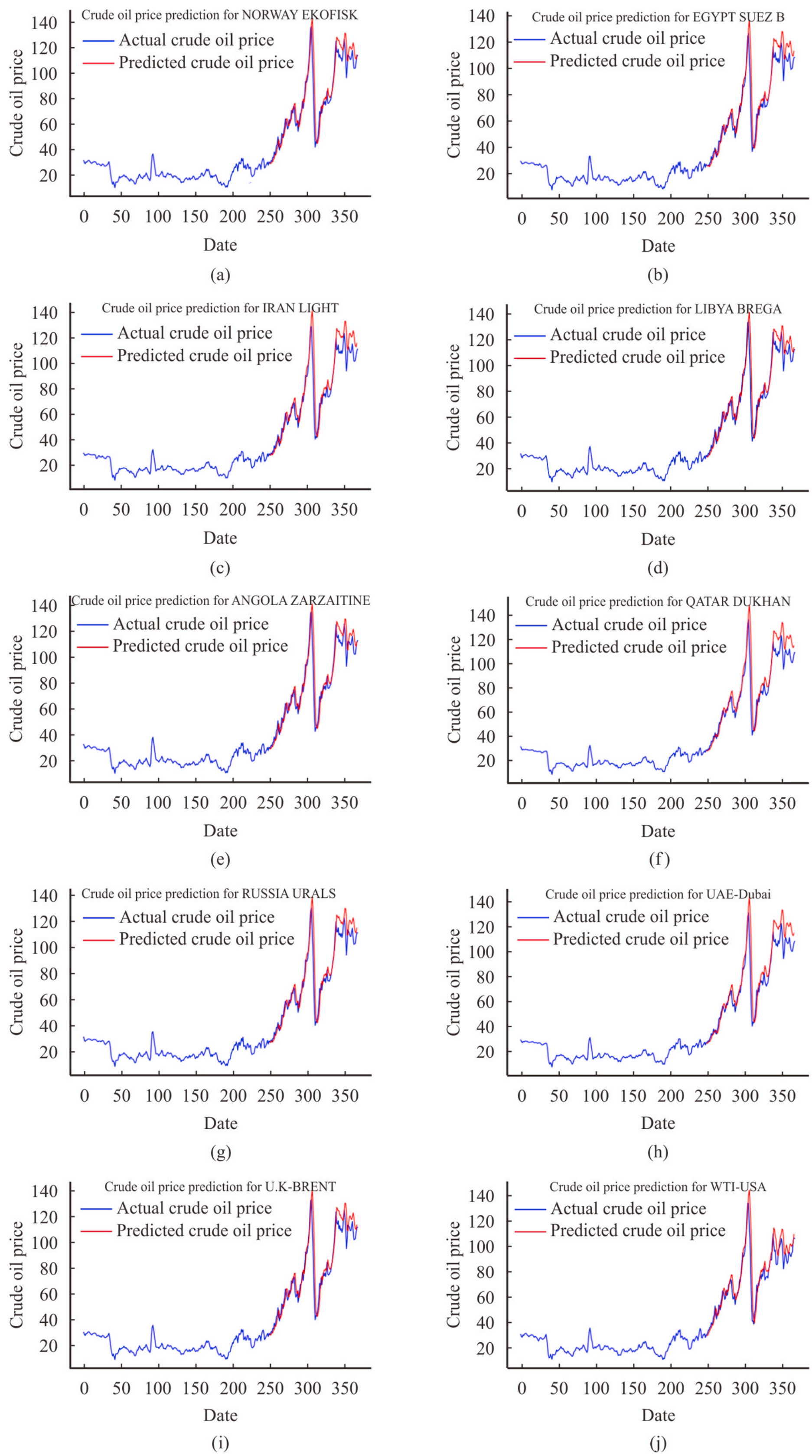

Fig. 7. (a)–(j). Ten (10) different crude oil prices across the world.

be higher than the one which is located at the periphery.”

Formally this can be stated as: Given an undirected simple network G, initially, the coreness c_i of every isolated node v_i (i.e $k_i = 0$) is defined as $c_i = 0$ and these nodes are removed before the k-core decomposition. Then in the first step of k-core decomposition, we first initialize k to 1, all vertex with degree $k = 1$ will be extracted. This will

Table 8
RMSE obtained while running as the experiment on other crude oil prices models.

Country/Type	RMSE
USA WTI	2.08
UK Brent	2.18
UAE Dubai	2.08
Russia Urals	2.20
Qatar Dukhan	2.07
Angola Zarzaitine	2.23
Libya Brega	2.21
Iran Light	2.02
Egypt Suez B	2.19
Norway Ekofisk	2.21

Table 9
Prediction Accuracy compared to other.

Models	RMSE
RW	6.832
MRS	5.616
SVR	5.867
SVR-B	5.872
FNN	5.428
FNN-B	5.079
SDAE	5.047
SDAE-B	4.995
LSTM	**2.91**

decrease the degree values of the remaining vertices. This procedure continually removes all the vertices whose residual degree $k \leq 1$, until all the remaining vertices' residual degrees $k > 1$. All nodes removed from first step of the decomposition constitute the 1-shell and their coreness ks equal one. Secondly, all the remaining vertices whose degrees $k = 2$ will be extracted in the first place. Then repeatedly extract all vertices whose residual degrees $k \leq 2$ until all the remaining vertices whose residual degrees $k > 2$. All the nodes removed in the second step of the decomposition constitute the 2-shell and their coreness ks equals two. This procedure continues until all vertices are extracted. Lastly, the coreness of a vertex vi equals its corresponding shell layer. Fig. 3 illustrates the k-core decomposition. Apparently, a vertex with a larger coreness signifies that the vertex is located in a more central position and is possibly more important in the network.

3.4. Transforming the datasets making it suitable for LSTM

3.4.1. Convert array of values into data matrix

With time series data, the sequence of values is important. A simple method that we can use is to split the ordered dataset into train and test datasets. The function takes two arguments: the dataset, which is a NumPy array that we want to convert into a dataset, and the look_back, which is the number of previous time steps to use as input variables to predict the next time period — in this case defaulted to 1."This default will create a dataset where X is the price at a given time (t) and Y is the price at the next time (t + 1)."Therefore, using the data obtained from our network we have 14.29, 15.00, 20.17, 17.44, 12.62. Comparing the Table 4 below with the dataset obtained we see that X = t and Y = t+1 and then transform it to stationery as obtained in Table 5.

3.4.2. Transform time series into stationary

Table 5 shows the stationary table obtained from Table 4.

3.4.3. Normalized the data

LSTMs are sensitive to the scale of the input data, specifically when the sigmoid (default) or tanh activation functions are used. It can be a good practice to rescale the data to the range of 0-to-1, also called normalizing Min-Max scaler is to adjust data into 0 to 1 by using min and max of the data as obtained in Table 6 from our example. MinMaxScaler in scikit-learn can use the Eq. (1) to transform datasets."

$$Xnew = \frac{X - Xmin}{Xmax - Xmin} \tag{1}$$

3.5. Fitting a stateful LSTM network model to the training data

LSTM network can learn and remember over a long sequence and does not rely on predefined window lagged observations as input. By default, LSTM layer in keras maintains state between data within one batch. The LSTM network expects the input data (x) in form (samples, time steps, features) and our data is of the form (samples, features), therefore we are framing the problem as one-time step for each problem. With a little trial and error, we can train the model and determine the best batch, size, epochs, and blocks or neurons.""

Batch Size: 15″
Epochs: 150″
Neurons: 4″

3.6. Evaluating the LSTM model on the test data

Once the LSTM is fit to the training data, it can be used to make prediction. We can fit the model once on all of the training data, and then predict each new time step one at a time from the test data. This is done by calling a *predict()* function on the model."

3.7. Accuracy of model evaluation

For the purpose of analyzing the prediction accuracy of the model, we employ root mean square error (RMSE) in Eq. (2) to evaluate the predictive performance of the model."

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(s_t - o_t)^2} \tag{2}$$

Where s_t, o_t are the actual price and the predictive value at time t."

4. Experimental results

The experiment in this paper was implemented in Python programming environment version 2.7.14 and version 3.7.1. The system requirements used in this experiment, windows 10, 64-bit operating system with 4 GB of RAM, Dual Core 2.0gz, 64 MB graphics card and at least 40 GB of hard disk space. The experiment was carried out using epochs of 150 and batch size of 15 to arrive at our result. The data we used for the experiment was obtained from www.opec.org/library and www.cbn.gov.ng/rates/crudeoil.asp. Table 7 presents the list of the datasets. We used these sets because they are used commonly as benchmark for petroleum exporting countries of which Nigeria is a member country. The datasets as divided into two contains 68% training set and 33% test set. A more elaborate description of the used sets is as follows:

(1) The Organization of the Petroleum Exporting Countries (OPEC) Secretariat is the executive arm of OPEC responsible for numerous publications and to disseminate data and information about OPEC member countries and oil industry in general.
(2) Central Bank of Nigeria (CBN) Data & Statistic is among the agencies having reliable dataset for collection and for research purposes; it is one of the most used datasets for accuracy.

4.1. Overall results

First, we check to see the volatility behavior of the dataset before conducting the experiment, the result is illustrated in Fig. 4. Fig. 5 illustrate network of datasets mapped on a network.

Fig. 6 shows the predicted and the actual crude oil price.

4.2. Results comparison

In order to evaluate the accuracy and effectiveness of our model, we carried out the experiment on ten (10) different crude oil prices across the world used by other researchers and the result is displayed in Fig. 7.

The average performance of the results in Table 8 shows the RMSE obtained across all the datasets which indicate that all the values obtained are less than 2.91 obtained from our main model and Table 9 shows the performance of our model in comparison to other prediction models with an accuracy of **2.91** which is an improvement over other model. This means that on both datasets our proposed technique has been able to increase accuracy **58%** as compared to best results achieved by the other techniques as shown in Eq. (3):

$$(2.91 / 4.995) * 100\% \approx 58.3\% \tag{3}$$

5. Conclusions and future work

Determining effective and efficient approach in predicting highly complex and volatile price like crude oil is a critical and challenging task in an economy of a nation. Most of the prediction techniques are designed focusing on statistical and econometrics point of view which has been helpful in numerous scenarios, however prediction using powerful AI tool like the LSTM of the DL is very rare. In this paper, we proposed a new crude oil price prediction technique based on complex network analysis and LSTM. In order to evaluate the effectiveness and robustness of the technique, we conducted the experiment on ten (10) different prices of crude oil across the world used by other researchers.

From the experiment conducted we can conclude that, during the training process, the selection of batch size and number of LSTM layers has a great influence on the objective function value, fitting effect, and running time. The appropriate batch size and number of LSTM layers can effectively improve the model.

Compared with the traditional and classic econometric prediction method, the model selects more datasets over a longer period of time as training samples. The LSTM prediction model has higher precision and wider application scenarios. The LSTM model can clearly predict the trend of crude oil price in the next time period.

The following points can be considered for future work:

(1) This paper only considers crude oil price in Nigeria, without necessary considering other factors such as, financial market, economic growth, dollar exchange rate, demand and supply etc. The model proposed in this thesis in build based on monthly data, which restrict the prediction horizons to months.
(2) The proposed technique can be extended by considering other factors that affect crude oil price volatilities such as, financial market, economic growth, exchange rate, demand and supply and the weather. And the horizon of the prediction can be widened by considering daily data.
(3) The proposed technique can be implemented with different dataset such as the stock market data in the future to further check the validity of the proposed technique.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.petlm.2019.11.009.

References

[1] Alley A. Asekomeh, H. Mobolaji, Y.A. Adeniran, Oil Price shocks and the Nigerian economic growth, Eur. Sci. J. 10 (2014) 375–391.
[2] J. Zhang, Y. Zhang, L. Zhang, A novel hybrid method for crude oil price prediction, J. Energy Econ. (2015) 1–11, https://doi.org/10.1016/j.eneco.2015.02.018.
[3] C.O. Augustine, Exchange rate fluctuations, oil prices and economic performance: empirical evidence from Nigeria, Int. J. Energy Econ. Policy 5 (2015) 502–506.
[4] A. Shabri, R. Samsudin, Daily crude oil forecasting using hybridizing wavelength and artificial neural network model, Math. Probl. Eng. (2014) 1–10, https://doi.org/10.1155/2014/201402.
[5] Y. Zhao, J. Li, L. Yu, A deep learning assemble approach for crude oil price forecasting, J. Energy Econ. (2017) 1–18, https://doi.org/10.1016/j.eneco.2017.05.023.
[6] H. Chiroma, S. Abdulkareem, T. Herawan, Evolutionary neural network model for west Texas intermediate crude oil price prediction, Appl. Energy (2015) 266–273.
[7] M.M. Mostafa, A.A. El-masry, Oil price forecasting using gene expression programming and artificial neural networks, J. Econ. Model. 54 (2016) 40–53, https://doi.org/10.1016/j.econmod.2015.12.014.
[8] M.R. Mahdiani, E. Khamehchi, A modified neural network model for predicting the crude oil price, Journal of Intellectual Economics 10 (2016) 71–77.
[9] R. Tehrani, F. Khodayar, A hybrid optimized artificial intelligent model to forecast crude oil using genetic algorithm, Afr. J. Bus. Manag. 5 (2011) 13130–13135, https://doi.org/10.5897/AJBM11.304.
[10] H. Shin, T. Hou, K. Park, C. Park, S. Choi, Prediction of movement direction in crude oil price based on semi-supervised learning, Decis. Support Syst. 55 (2013) 348–358, https://doi.org/10.1016/j.dss.2012.11.009.
[11] L. Yu, W. Dai, L. Tang, A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting, J. Eng. Appl. Artif. Intell. (2016) 1–12, https://doi.org/10.1016/j.engappai.2015.04.016i.
[12] Y. Chen, K. He, K.G.F. Tso, Forecasting crude oil price: a deep learning-based model. International conference on information technology and quality management, Procedia Comput. Sci. 122 (2017) 300–307, https://doi.org/10.1016/j.procs.2017.11.373.
[13] H. Wang, F. Zhang, X. Xie, M. Guo, Deep knowledge-aware network for news Recommendation, ACM 2 (2018) 1–10.
[14] X. Ma, Y. Jin, Q. Dong, A generalized dynamic fuzzy neural network based on singular spectrum analysis optimized by brain storm optimization for short-term wind speed forecasting, Appl. Soft Comput. 54 (2017) 296–312.
[15] J. Wang, X. Li, T. Hong, S. Wang, A semi-heterogeneous approach to combining crude oil price forecasts, J. Inf. Sci. (2018) 279–292.
[16] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, The MIT Press, Cambridge MA, 2016.
[17] A. Prieto, B. Prieto, E. Martinez-Ortigosa, et al., Neural networks: an overview of early research, current frameworks and new challenges, Neurocomputing 214 (2016) 242–268.
[18] J. Patterson, A. Gibson, Deep Learning a Practitioners Approach, O'Reilly Media,Inc, Newton, MA, 2017.
[19] V. Sze, Y. Chen, T. Yang, J.S. Emer, Efficient processing of deep neural networks, A tutorial and survey, CoRRabs/1703.09039, 2017. http://arxiv.org/abs/1703.09039.
[20] W. Liu, Z.X.L. Wang, et al., A survey of deep neural network architectures and their applications, Neurocomputing 234 (2017) 11–26.
[21] Y. Ding, A novel decompose-assemble methodology with AIC-ANN approach for crude oil forecasting, J. Energy 154 (2018) 328–336.
[22] H. Naser, Estimating and forecasting the real prices of crude oil: a data rich model using a dynamic model averaging (DMA) approach, Energy Econ. 56 (2016) 75–87.
[23] L. Fan, S. Pan, Z. Li, H. Li, An ica-based support vector regression scheme for forecasting crude oil prices, Technol. Forecast. Soc. Chang. 112 (2016) 245–253.
[24] O. Saltik, S. Degirmen, M. Ural, Volatility modelling in crude oil and natural gas price, Istanbul Conference of Economics and Finance, ICEF 2015, 22-23 October 2015, Istanbul, Turkey, 38 2015, pp. 476–491.
[25] F. Chiu, C. Hsu, A. Ho, C. Chen, Modeling the price relationships between crude oil, energy crops and biofuels, Energy 109 (2016) 845–857.
[26] A. Jedidzadeh, A. Serletis, The global crude oil market and biofuel agricultural commodity prices, J. Econ. Asymmetries 18 (2018) 1–9.
[27] A. Khashman, N. Nwulu, Intelligent Prediction of Crude Oil Price Using Support Vector Machines, Applied Machine Intelligence and Informatics (SAMI), 2011, IEEE 9th international symposium on IEEE, 2011, pp. 165–169.
[28] S. Li, Y. Ge, Crude oil price prediction based on dynamic correcting support vector regression machine, Abstr. Appl. Anal. (2013) 666–686.
[29] X. Li, I. Yu, I. Tang, W. Dai, Coupling Firefly Algorithm and Least Squares Support Vector Regression for Crude Oil Price Forecasting, Business Intelligence and Financial Engineering (BIFE), 2013 Sixth International Conference on IEEE, (2013), pp. 80–83.
[30] S. Gao, Y. Lei, A new approach for crude oil price prediction based on stream learning, J. Geosci. Front. 8 (2017) 183–187, https://doi.org/10.1016/j.gsf.2016.08.002.
[31] A.A. Godarzi, R.M. Amiri, A. Talaei, T. Jamasb, Predicting oil price movements: a dynamic neural network Approach, Energy Policy 68 (2014) 371–382.
[32] L. Yu, Y. Zhao, L. Tang, A compressed based AI learning paradigm for crude oil price forecasting, J. Energy Econ. 46 (2014) 236–245.
[33] L.A. Gabralla, R. Jammazi, A. Abraham, Oil Price Prediction Using Ensemble Machine Learning, 2013 International Conference on Computing, Electrical And Electronic Engineering (ICCEEE), 2013, pp. 674–679.

[34] R. Jammazi, C. Alou, Crude oil price forecasting: experimental evidence from wavelet decomposition and neural network modeling, Energy Econ. 34 (2012) 828–841.

[35] T. Xion, Y. Bao, Z. Hu, Beyond one-step-ahead forecasting: evaluation of alternative multi-step-ahead forecasting models for crude oil prices, Energy Econ. 40 (2013) 405–415.

[36] Q. Ji, H. Zhan, Y. Fan, Identification of global oil trade patterns: an empirical research based on complex network theory, J. Energy Conv. Manag. (2014) 1–10, https://doi.org/10.1016/j.enconman.2013.12.072.

[37] Y. Zou, M. Small, Z. Liu, J. Kurths, Complex network approach to characterized the statistical features of the sunspot series, New J. Phys. 46 (2014) 236–245, https://doi.org/10.1016/j.eneco.2014.09.019.

[38] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J.C. Nuño, From time series to complex networks: the visibility graph, Proc. Natl. Acad. Sci. U. S. A 105 (2008) 4972-497.

[39] N. Wang, D. Li, Q. Wang, Visibility graph analysis on quarterly macroeconomic series of China based on complex network theory, Physica A 391 (2012) 6543–6555.

[40] A.M. Nuñez, L. Lacasa, J.P. Gomez, B. Luque, Visibility algorithms: a short review, in: Y. Zhang (Ed.), New Frontiers in Graph Theory, InTech, Rijeka, Croatia, 2012, pp. 119–152.

[41] M. Kitsak, L.K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H.E. Stanley, et al., Identification of influential spreaders in complex networks, Nat. Phys. 6 (2010) 888–893.