Considering interpretability of the LSTM Architecture for Oil Stocks Prices Prediction

Javad T. Firouzjaee
Department of Physics, K. N. Toosi University of Technology,
P. O. Box 15875-4416, Tehran, Iran
PDAT Laboratory, Department of Physics,
K. N. Toosi University of Technology,
P. O. Box 15875-4416, Tehran, Iran and
School of Physics, Institute for Research in Fundamental Sciences (IPM),
P. O. Box 19395-5531, Tehran, Iran

Pouria Khalilian
PDAT Laboratory, Department of Physics,
K. N. Toosi University of Technology,
P. O. Box 15875-4416, Tehran, Iran and
Qorpi AI Workgroup, Innovation center,
K. N. Toosi University of Technology,
P. O. Box 19697-64499, Tehran, Iran

Oil companies are among the largest companies in the world whose economic indicators in the global stock market have a great impact on the world economy and market due to their relation to gold, crude oil, and the dollar. To quantify these relationships, we use correlation features and relationships between stocks with the dollar, crude oil, gold, and stock indices of major oil companies, create data sets, and perform continuous and discrete correlation analyses with each other. To predict the stocks of different companies, we use Recurrent Neural Networks (RNNs) and LSTM, because these stocks change in time series. We carry out empirical experiments and perform on the stock indices dataset to evaluate the prediction performance in terms of several common error metrics such as Mean Square Error (MSE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). The received results are promising and present a reasonably accurate prediction for the price of oil companies’ stocks in the near future. Despite the volatility of the investigated systems in continuous and discrete correlation analysis, LSTM has a high interpretability ability to investigate surprising.

*Electronic address: firouzjaee@kntu.ac.ir
†Electronic address: pouriya@email.kntu.ac.ir
Machine Learning, which is one of the subfields of Artificial Intelligence, has its applications in various fields including Economics, Medicine [1], Cosmology [2], Particle physics [3], Robotics [4], etc. The machine learns and models based on non-explicit programming based on the datasets that we have collected in the preprocessed datasets, and we compare the modeled data with the real data. Thus, we can see the data extent accurately which is modeled by the machine. Artificial Neural Networks are derived from Natural Neural Networks in living things, which are a subset of Machine Learning, designed to predict responses from complex systems. One of the most famous neural networks is Recurrent Neural Networks or RNNs that function close to the human brain.

We know that the largest market in the field of Energy belongs to the oil companies. In the field of oil, there are large companies around the world that have a very high impact. In the world economy, oil can be considered the most vital factor of the economy, because, for example, if the export or import of oil from many countries is sanctioned, the economy of that country will be practically paralyzed, especially for countries with Oil-dependent economies, eg. Persian Gulf countries. Stock indices of Oil companies are among the most important indicators in the global stock market, which have correlations between the shares of oil companies, Gold, US dollar, crude oil, which has an impact on them.

Although Neural Networks are not interpretive and can not interpret these correlations well, they can be used in learning and modeling. The LSTM is one of the most powerful architectures in Recurrent Neural Networks that has solved the problem of gradient vanishing in Recurrent Neural Networks and can help us to predict the stock market more accurately. The oil stock market is like an unstable dynamic system with many non-linear correlations, and researchers have proposed various methods to predict the behavior of this system.

To review the literature oil research path, we can mention Alvarez-Ramirez et al. work [5] which analyzed the auto-correlations of international crude oil. After carrying out several tests, in 2005, Moshiri and Foroutan [6] concluded that oil stock markets have a recursive architecture because they are time series. They used three methods ANN, GARCH, and ARMA, and the best results come from the ANN method.

Author in [7] published an article using Multilayer Neural Networks, which examined the relationship between crude oil prices and current prices. Moreover, they showed that the future prices of crude oil contain new information about oil spot price detection. Ye et al. [8] studied the changing relationships between the prices of crude oil and several other factors from January 1992 to December 2007 by the Short-Run Crude Oil Price Forecast Model.

Chen and colleagues developed a model based on deep learning and used this model to model the unknown non-linear behavior of WTI stocks [10]. QI, KHUSHI, and POON
used different recursive neural network architectures including LSTM, GRU, BiLSTM, RNN to model the Forex market and obtained significant results from these models to predict several currency pairs. They used a database that relates to ELLIOT method information, one of the stock market forecasting methods [11].

In 2018, Gupta and Pandey predicted crude oil prices by using LSTM network [12], and following that Cen and Wang applied deep learning algorithms to anticipate the volatility behavior of crude oil prices [13]. To solve the chaotic and nonlinear features of crude oil time series Altan and Karasu used a new crude oil price prediction model is proposed in this study, which includes the long short-term memory (LSTM), technical indicators [15].

In 2017, Arfaoui and Rejeb published an article examining the effects and relationships of stock markets, oil, dollars, and gold based on global market documentation. They concluded that oil prices are significantly affected by stock markets, gold, and the dollar and that there are always indirect effects, which also confirms the presence of correlations in the global market [16]. In this paper, we compare this correlation feature and the relationships between stocks with the dollar, crude oil, gold, and major oil company stock indices, we create datasets and compare the results of forecasts with real data.

The paper is organized as follows: Section II is dedicated to an analysis of the correlation between different shares and economic stocks. In Section III, we apply LSTM architecture to predict oil prices. Finally, we conclude and summarize the main results in Section VI.

II. OIL SHARES AND CORRELATION

The relation between oil companies’ stocks on one hand and the prices of other assets like gold, crude oil, and the price of the US dollar, on the other hand, is intrinsically interesting for many economical and political reasons. Before calculating the correlation coefficient, it has worth having a short introduction for oil companies.

A. Oil companies

In the global economy, large companies operating in the field of energy and oil are among the most authoritative companies in the global economy. Companies such as Total, BP, Cairn Energy, Schlumberger are among the most authoritative companies. These companies are trading in energy and oil and their investors and shareholders are among the largest investors and shareholders in the world, for whom the trend of changing the indicators of these companies in the stock market is very worthful. Moreover, the strong dependence of industry in various fields on oil has made oil a major factor influencing the
policies of each country. In the following, we introduce some of these companies which we analyze their shares in this paper.

WTI: One of known light sweet crude oil is West Texas Intermediate (WTI) which refers as one of the main global oil benchmarks. WTI has high quality is easy to refine and is sourced primarily from inland Texas.

BP.L: BP plc is a British multinational oil and gas company based in London, England. It is one of the world’s seven oil and gas companies.

FP.PA: TotalEnergies SE is a French multinational unified oil and gas company which is established in 1924 and is one of the seven major oil companies.

SLB.PA: Schlumberger Limited is an oil field services company whose member have more than 140 nationalities working in more than 120 countries. Schlumberger executive offices are located in Paris, Houston, London, and The Hague.

B. Correlation between shares and economic stocks

The crude oil index is the most important characteristic in determining the index of oil companies. The sale of crude oil in the world market is done in dollars. Commonly, the dollar can change the value of the indexes of oil companies. Moreover, the gold value can cause changes and fluctuations in currency indices. Their effects on each others can be obtained by calculating the correlation coefficient between their data [16]. In this article, the value of the stock indices that we have examined, from 08/03/2009 to 07/01/2021, the results of the correlation calculation are shown in the table (I):

|       | USD          | WTI          | GOLD         | BP           | TOTAL        | Schlumberger | Cairn Energy |
|-------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| USD   | 1.000000     | -0.357051    | -0.002025    | -0.214303    | -0.088397    | 0.225169     | 0.046451     |
| WTI   | -0.357051    | 1.000000     | -0.046961    | 0.371175     | 0.355021     | 0.377098     | 0.490978     |
| GOLD  | -0.002025    | -0.046961    | 1.000000     | -0.464169    | -0.553079    | -0.514969    | -0.141593    |
| BP    | -0.214303    | 0.371175     | -0.464169    | 1.000000     | 0.871033     | 0.381587     | 0.521461     |
| TOTAL | -0.088397    | 0.355021     | -0.553079    | 0.871033     | 1.000000     | 0.529159     | 0.494148     |
| Schlumberger | 0.225169 | 0.377098 | -0.514969 | 0.381587 | 0.529159 | 1.000000 | 0.431568 |
| Cairn Energy | 0.046451 | 0.490978 | -0.141593 | 0.521461 | 0.494148 | 0.431568 | 1.000000 |

This table is represented as an array where its principal diameter is equal to one, and the other component represents the correlation coefficient between the value of stock
indices introduced in the global market.

To have a better presentation, we draw the Heatmap diagram which is a type of information visualization (data visualization) in which the value of each cell of the matrix input is displayed in one color. The correlation coefficient obtained in the previous part is shown in the diagram of the thermal map below Fig. 1. The lighter the color, the more direct the relationship and the higher the direct correlation, and the darker color, the higher the inverse relationship and correlation.

![Heatmap of the correlation coefficients in the Table (I).](image)

**FIG. 1:** Heatmap of the correlation coefficients in the Table (I).

**C. Discrete correlation analysis for oil shares**

For a more detailed analysis of the relationships between WTI, USD, and GOLD with shares of TOTAL, BP, Schlumberger, and Cairn Energy oil companies, we discretized the data and then checked the correlations. From 2013/6/24 to 2021/4/7, we discretized about 2000 data for each stock into 50 40-day data sets. Discrete correlations of the BP company data set with WTI, USD and GOLD data sets are shown in the Fig. 2.
Correlation histograms of BP stock index with WTI, USD and GOLD stock index are shown in Fig. (3). Histograms have four columns according to the values of correlation analysis, which are shown in Table (II).

The statistical parameters of average, median, variance and standard deviation for the correlations of the BP stock index with WTI, USD and GOLD are shown in Table (III) and Fig. (4):

Discrete correlations of the Total company data set with WTI, USD and GOLD data sets are shown in the Fig. (5).
TABLE II: Correlations amount.

|                  | Corr(-1,-0.5) | Corr(-0.5,0) | Corr(0,0.5) | Corr(0.5,1) |
|------------------|---------------|--------------|-------------|-------------|
| Count of corr(BP-WTI) | 1.0           | 13.0         | 15.0        | 20.0        |
| Percent of corr(BP-WTI)  | 2.0           | 26.0         | 30.0        | 40.0        |
| Count of corr(BP-USD)   | 6.0           | 12.0         | 16.0        | 15.0        |
| Percent of corr(BP-USD)  | 12.0          | 24.0         | 32.0        | 30.0        |
| Count of corr(BP-GOLD)  | 15.0          | 8.0          | 15.0        | 11.0        |
| Percent of corr(BP-GOLD) | 30.0          | 16.0         | 30.0        | 22.0        |

TABLE III: BP discrete correlations statical parameters.

|                  | Median     | Mean       | Variance   | Standard deviation |
|------------------|------------|------------|------------|--------------------|
| USD-BP           | 0.207317   | 0.142085   | 0.221153   | 0.470269           |
| GOLD-BP          | 0.119667   | -0.029858  | 0.312736   | 0.559228           |
| WTI-BP           | 0.325000   | 0.274804   | 0.209705   | 0.457936           |

FIG. 4: BP discrete correlations statical parameters analysis heatmap.

Correlation histograms of Total stock index with WTI, USD and GOLD stock index are shown in Fig. (6). Histograms have four columns according to the values of correlation.
FIG. 5: Total discrete correlations analysis.

analysis which are shown in Table IV.

FIG. 6: Total discrete correlations histograms.

TABLE IV: Correlations amount of Total.

| Count of corr(Total-WTI) | Corr(-1,-0.5) | Corr(-0.5,0) | Corr(0,0.5) | Corr(0.5,1) |
|--------------------------|---------------|--------------|-------------|-------------|
| Count of corr(Total-USD) | 6.0           | 34.0         | 30.0        | 30.0        |
| Percent of corr(Total-USD)| 14.0          | 22.0         | 20.0        | 20.0        |
| Count of corr(Total-GOLD)| 10.0          | 15.0         | 15.0        | 9.0         |
| Percent of corr(Total-GOLD)| 22.0          | 30.0         | 30.0        | 18.0        |
The statistical parameters of average, median, variance, and standard deviation for the correlations of the Total stock index with WTI, USD and GOLD are shown in Table (V) and Fig. (7):

### TABLE V: Total discrete correlations statistical parameters

|                | Median | Mean   | Variance | Standard deviation |
|----------------|--------|--------|----------|--------------------|
| USD-Total      | 0.064403 | 0.073098 | 0.233880 | 0.483611           |
| GOLD-Total     | -0.011213 | -0.021289 | 0.270061 | 0.519674           |
| WTI-Total      | 0.164420  | 0.164086  | 0.195751  | 0.442437           |

FIG. 7: Total discrete correlations statistical parameters analysis heatmap.

Discrete correlations of the Schlumberger company data set with WTI, USD and GOLD data sets are shown in the Fig. (8).

Correlation histograms of Schlumberger stock index with WTI, USD and GOLD stock index are shown in Fig. (9). Histograms have four columns according to the values of correlation analysis which are shown in Table (VI).
The statistical parameters of average, median, variance and standard deviation for the correlations of the Schlumberger stock index with WTI, USD and GOLD are shown in Table (VII) and Fig. (10):

|            | Corr(-1,-0.5) | Corr(-0.5,0) | Corr(0,0.5) | Corr(0.5,1) |
|------------|---------------|--------------|-------------|-------------|
| Count of corr(Schlumberger-WTI) | 3.0           | 12.0         | 21.0        | 13.0        |
| Percent of corr(Schlumberger-WTI) | 6.0           | 24.0         | 42.0        | 28.0        |
| Count of corr(Schlumberger-USD)   | 3.0           | 19.0         | 12.0        | 15.0        |
| Percent of corr(Schlumberger-USD) | 6.0           | 38.0         | 26.0        | 30.0        |
| Count of corr(Schlumberger-GOLD)  | 11.0          | 13.0         | 12.0        | 13.0        |
| Percent of corr(Schlumberger-GOLD)| 22.0          | 26.0         | 26.0        | 26.0        |
TABLE VII: Schlumberger discrete correlations statistical parameters.

|                      | Median  | Mean    | Variance | Standard deviation |
|----------------------|---------|---------|----------|--------------------|
| USD-Schlumberger     | 0.043825| 0.100825| 0.224941 | 0.474280           |
| GOLD-Schlumberger    | 0.049048| 0.015501| 0.309095 | 0.555963           |
| WTI-Schlumberger     | 0.299134| 0.193702| 0.175441 | 0.418857           |

FIG. 10: Schlumberger discrete correlations statistical parameters analysis heatmap.

Discrete correlations of the Cairn Energy company data set with WTI, USD and GOLD data sets are shown in the FIG. (11).

Correlation histograms of Cairn Energy stock index with WTI, USD and GOLD stock index are shown in Fig. (12). Histograms have 4 columns according to the values of correlation analysis, which are shown in Table (VIII).

The statistical parameters of average, median, variance and standard deviation for the correlations of the Cairn Energy stock index with WTI, USD and GOLD are shown in
TABLE VIII: Correlations amount of Cairn Energy.

|                     | Corr(-1,-0.5) | Corr(-0.5,0) | Corr(0,0.5) | Corr(0.5,1) |
|---------------------|---------------|--------------|-------------|-------------|
| Count of corr(Cairn Energy-WTI) | 4.0           | 9.0          | 18.0        | 19.0        |
| Percent of corr(Cairn Energy-WTI)  | 8.0           | 18.0         | 36.0        | 38.0        |
| Count of corr(Cairn Energy-USD)   | 4.0           | 11.0         | 17.0        | 18.0        |
| Percent of corr(Cairn Energy-USD)  | 8.0           | 22.0         | 34.0        | 36.0        |
| Count of corr(Cairn Energy-GOLD)  | 12.0          | 12.0         | 14.0        | 12.0        |
| Percent of corr(Cairn Energy-GOLD) | 24.0          | 24.0         | 28.0        | 24.0        |

Table IX and Fig. 13:
TABLE IX: Cairn Energy discrete correlations statically parameters.

|                      | Median | Mean  | Variance | Standard deviation |
|----------------------|--------|-------|----------|--------------------|
| USD-Cairn Energy     | 0.248828 | 0.198628 | 0.217601 | 0.46647            |
| GOLD-Cairn Energy    | 0.016285 | -0.008477 | 0.290874 | 0.539328           |
| WTI-Cairn Energy     | 0.321247 | 0.261865 | 0.214425 | 0.463061           |

FIG. 13: Cairn Energy discrete correlations statistical parameters analysis heatmap.

III. LSTM ARCHITECTURE FOR OIL PRICES

Long Short Term Memory, LSTM is an architecture for RNNs to avoid the gradient vanishing problem in RNNs in memory building. LSTM architecture has three gates, forget gate, input gate, and output gate as shown in Fig. 14. LSTM can read, write and delete information from its memory. The LSTM structure has the capability to identify which cells are stimulated and compressed based on the previous state, available memory, and current input.

Forget gate determine what information data we’re going to throw away from the cell state. In this gate, information is deleted or stored based on the output of the sigmoid
function. This gate cell is formulated as follows:

$$f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f).$$  \hspace{1cm} (1)$$

Input gate determine what new data we’re going to store in the cell state. The input gate is actually a gate for writing memory which is represented as

$$i_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f),$$ \hspace{1cm} (2)

$$\tilde{c}_t = \tanh(W_c \times [h_{t-1}, x_t] + b_c).$$ \hspace{1cm} (3)

Cell state drops the information about past subjects and adds new information as follows

$$c_t = f_t \times c_{t-1} + \tilde{c}_t \times i_t.$$

Output gate decide what we’re going to output and reading from Memory which is formulated as

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t \times \tanh(c_t).$$ \hspace{1cm} (4)

A. Data and Features

In this paper, we use four major (open, high, low, and closing prices) Oil stock pairings- FP.PA (Total), CNE.L (Cairn Energy), BP.L (BP), SLB.PA (Schlumberger)
with their daily interval data and use for comparison stock pairings - WTI, gold, dollar with their daily interval data from 2009/8/14 to Sep 2020/7/19 for training and data from 2020/7/20 to 2021/7/15 for testing. Data was extracted from Alpha Vantage API.

\section*{B. Hyper-parameter selection}

The fundamental determination of hyper-parameters for the neural network has a very important effect on the modeling results. History point is one of the most important hyper-parameters that has a very important effect on the results in modeling with recurrent neural networks. Basic determination of this value prevents possible problems in modeling, including gradient vanishing. To determine this amount, various correlations of the close price of each stock, which is the purpose of prediction, have been taken with its own future data, to be able to see how much past prices affect future prices. Basically, in order to be able to determine the amount of memory examined for modeling with recurrent neural networks, we use the history point. The results of the correlations of the close price of each stock from 2013/6/23 to 2017/6/12 with each of the dates in Fig. (15), Fig. (16), Fig. (17) and Fig. (18) are as follows:

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{bp_history.png}
\caption{BP past and future correlations.}
\end{figure}

Due to the fact that in the (II C) section, we have discretized the data as 40-day packages and then performed the discrete correlation analysis, and according to Fig. (15), (16), (17), and (18), we can see that the data has a strong direct relationship with the future 40 days. According to the obtained results, we set the value of the history point equal to 40, so that we can check the interpretability of LSTM for predicting the data of oil stocks.
based on the results obtained from the modeling and the results of the (II C) section.
Adam’s\cite{17} optimization function is used with 0.0005 learning rate, 10 percent data is validation data, epochs value is equal to 50 and batch size is equal to 25. For all simulations with LSTM, these values are the same.

C. Evaluation Measures

To measure the badness of the model, people use Mean Square Error (MSE), Root means square error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), as defined below: For measuring the average of the squares of the errors, MSE measure can be used which is defined as the following formula:

\[ MSE = \frac{\sum_{i=1}^{n} (true - Prediction)^2}{n}. \]

To measure the average root of the squares of the errors we use RMSE which is defined as:

\[ RMSE = \sqrt{\frac{\sum_{i=1}^{n} (true - Prediction)^2}{n}}, \]

which describes how spread out these residuals are by the standard deviation of prediction errors. To measure the average of absolute error we use MAE which calculates errors between pair observations. It is defined as:

\[ MAE = \frac{\sum_{i=1}^{n} |true - Prediction|}{n}. \]

Finally, to measure the size of the error in percentage terms we use MAPE which is calculated for the mean of the absolute percentage errors of prediction. It is defined as the following formula:

\[ MAPE = \frac{\sum_{i=1}^{n} |true - Prediction|}{\sum_{i=1}^{n} true} \times 100. \]  \hspace{1cm} (5)

Now, we are in the position where we can apply these evaluation measures as error metrics in our prediction.

IV. DISCUSSION AND CONCLUSION

In this work, we highlighted the potential of deep learning new tools as an alternative or at least a complement to traditional forecasting methods.
First, we studied the correlation coefficient between oil companies’ stocks, USD, WTI crude oil, and gold. Then we put the stock data discretely and carried out the correlation analysis discretely. In Sec. (II B), we got that there is a correlation between the shares of the oil companies in question together with the gold, dollar, and crude oil indices. According to Table (I), the following results have been obtained:

- BP stock index has a weak direct relationship with WTI and this stock has a weak inverse relationship with USD and gold.

- BP stock index has a strong direct relationship with TOTAL, CAIRN ENERGY stock index and has a weak direct relationship with Schlumberger.

- TOTAL stock index has almost no correlation with USD, it has a weak direct relationship with the WTI index, and it has a strong inverse relationship with the gold index.

- TOTAL index shares have a direct relationship with Cairn Energy, Schlumberger, and BP indices.

- The Schlumberger stock index has a weak direct correlation with USD and WTI and also a strong inverse correlation with gold

- The Schlumberger stock index has a weak direct relationship with BP and Cairn Energy and a strong direct relationship with the Total stock index

- The Cairn Energy stock index has almost no relationship with the USD, it has a weak direct relationship with the WTI stock index, and it has a weak inverse relationship with the gold index.

- Cairn Energy index shares have a strong direct relationship with BP and also have a weak direct relationship with Cairn Energy and Schlumberger.

According to the interpretation of Table (I), we can conclude that the stock index of the four companies reviewed from 2013/6/24 to 2021/4/7 has a weak direct relationship with the WTI crude oil stock index. and they have an inverse relationship with gold. They have no correlation with the dollar or have a weak relationship and have a strong direct
relationship with each other.

Second, to check the results obtained from part (II B), how accurate they are and how stable they are before modeling and checking the results of modeling with the help of LSTM, we check the results of part (II C), which is a discrete analysis of correlations.

- According to Fig. (2), the correlations of BP shares with the index of gold, USD, and WTI stocks have a lot of fluctuations in the abnormal correlation intervals, and these fluctuations are not in a specific correlation interval. In Fig. (3) and Table (II), about 30% of the correlations show a weak direct relationship with WTI, which shows a contradiction with the results of Table (I). and 40% of the correlations show a strong direct relationship. 36% of the correlations show an inverse relationship with the USD, which contradicts Table (I), and about 64% of the correlations show a strong direct relationship with the USD. About 46% of the correlations show an inverse relationship with gold and about 54% of the correlations show a direct relationship with gold, which shows a contradiction with the results of Table (I). The average correlation of BP shares with the WTI stock index is 0.27, which indicates a weak direct relationship and is roughly consistent with the results of Table (I), but this average with a standard deviation of 0.45 indicates high dispersion. The noteworthy point is the median value, which is 0.32, which is between the result obtained from Table (I) and the average of the correlations, which indicates that the correlations that have an inverse relationship are more weighted. This can indicate the effects of indirect factors such as political events. The average correlations of BP company’s shares with the gold stock index almost represent the absence of relationship and anti-correlatedness, which is a contradiction with the results of table (I). The median has a smaller value than the average, which indicates that correlations have more weight than the reverse relationship between they evoke two stocks. The average correlations of this company’s shares with the dollar show a weak direct relationship, which is a contradiction with the results of Table (I), and the average of these correlations also shows the weight of the correlations that show a direct relationship.

- Correlations of shares of TOTAL, Schlumberger, and CNE companies, according to Fig. (5), Fig. (8), and Fig. (11) like BP company’s correlations shown in Fig. (2), have high fluctuations in correlation ranges. According to Fig. (6) and Table (IV), about 60% of the correlations show a direct relationship between TOTAL shares with the stock index and WTI, which shows a contradiction with the results of Table (I). About 56% of the correlations show a direct relationship between stocks TOTAL is with USD shares, which shows a contradiction with the results of Table (I), and also about gold, we can say that it confirms the results of Table (I).median correlations of TOTAL shares with USD shares and gold are about zero,
which is following the results of table (I) in the case of USD shares, but in the case of gold, it contradicts the results of table (I). Of course, this is an important point. According to the value of the standard deviation, there is a high dispersion between the correlation values. The average and median correlation of this stock with WTI is in accordance with the results of Table (I).

- According to Fig. (9) and Table (VI), about 70% of the correlations show a direct relationship between Schlumberger company shares and WTI shares, which is roughly in line with the results of Table (I). About 56% of the correlations show a direct relationship between Schlumberger shares and USD. About 52% of the correlations of these stocks with gold stocks show a contradiction with the results of Table (I). The mean correlations in table (IX) for Schlumberger stocks with WTI indicate a weak direct relationship. The mean and average of these shares with gold shares are equal to zero, which contradicts the results of Table (I), and in the case of the mean and median of these shares with dollars, it shows the greater weight of correlations that show a direct relationship.

- According to Fig. (12) and Table (VIII), about 72% of the correlations show a direct relationship between Cairn Energy shares and WTI, which is consistent with the results of Table (I). 70% of the correlations represent the direct relationship between these stocks and USD, which contradicts the results of Table (I), which indicated the absence of a relationship. 52% of correlations show a direct relationship between Cairn Energy shares and gold shares, which shows the contradiction of the result of Table (I). The median and average correlations of Cairn Energy shares with USD represent a weak direct relationship, which contradicts the results of Table (I). The mean and mean correlation of this sum with gold shows a contradiction with the result of Table (I) because it shows the absence of a relationship. Also, the mean and average correlations of these stocks with WTI, show a direct relationship, which is following the results of Table (I).

According to the results obtained from the discrete analysis of correlations, we can note that the result of correlations does not always remain stable and does not increase in a specific correlation interval, but fluctuates in different correlation intervals. In the oil markets, according to the influence of factors such as political decisions on the price, political changes and wars of solidarity change in time.

Third, to predict the stocks of different companies, we have used Recurrent Neural Networks with LSTM architecture, because these stocks change in time series. We carried out empirical experiments and perform on the stock indices dataset to evaluate the prediction performance in terms of several common error metrics MSE, MAE, RMSE, and MAPE. Let us summarize the results with these items:
• In the table (X), we saw that adding the WTI (crude oil), gold, and dollar indices do not improve the model prediction and do not reduce the badness of the model (error metrics), and the same four main features have the lowest amount of cost functions. We see BP empirical analysis result in Fig. (19):

| Feature  | MSE    | RMSE   | MAE    | MAPE   |
|----------|--------|--------|--------|--------|
| WTI      | 0.02634| 0.16232| 0.01026| 0.08321|
| Main Dataset | 0.00372| 0.06105| 0.00386| 0.08335|
| Dollar   | 0.13510| 0.36756| 0.02324| 0.08096|
| Gold     | 0.02252| 0.15008| 0.00949| 0.08311|

FIG. 19: (BP Real Prices vs Predicted Prices)

• In the table (XI), we saw that adding the crude oil index and the dollar index did not reduce the cost function, but at the same time, when the machine learned the gold index in the learning process, it was able to have better measurement and reduce costs significantly and help us improve modeling as shown in the Carien Energy diagrams in Fig. (20).

| Feature  | MSE    | RMSE   | MAE    | MAPE   |
|----------|--------|--------|--------|--------|
| WTI      | 146.4406| 12.1012| 0.7653 | 0.3544|
| Main Dataset | 78.5752| 8.8642 | 0.5606 | 0.4298|
| Dollar   | 79.4293| 8.9123 | 0.5636 | 0.4272|
| Gold     | 48.4487| 6.9605 | 0.4402 | 0.4733|

TABLE XI: Experiment Result For Carien Energy Shares
In the table (XII), we see that the WTI index has increased the value of the cost function and adding the gold and dollar indices has made almost no change in the learning process. We see the diagrams for the Schlumberger company in Fig. (21).

|               | MSE     | RMSE    | MAE     | MAPE   |
|---------------|---------|---------|---------|--------|
| WTI           | 0.00964 | 0.09818 | 0.00621 | 0.02830|
| Main Dataset  | 0.00102 | 0.03203 | 0.00202 | 0.02569|
| Dollar        | 0.00112 | 0.03361 | 0.03361 | 0.025234|
| Gold          | 0.00108 | 0.03296 | 0.00208 | 0.025690|

In the table (XIII) for Total Shares, we see that none of the indicators of WTI, dollar, gold has improved the learning process which can be seen in Fig. (22).

|               | MSE     | RMSE    | MAE     | MAPE   |
|---------------|---------|---------|---------|--------|
| WTI           | 0.000297| 0.01723 | 0.00109 | 0.02621|
| Main Dataset  | 2.9648e-07 | 0.0005442 | 3.4437e-05 | 0.02582|
| Dollar        | 2.3122e-05 | 0.00480 | 0.00030 | 0.02613|
| Gold          | 0.00040 | 0.02001 | 0.00126 | 0.02579|

Overall, the above results show that the range of errors for each stock with different modeling is in the same order. According to the results of part (II C) and modeling results, it can be concluded that LSTM has the necessary interpretability to predict oil stocks. One
of the main applications of examining different correlations and obtaining indirect factors to improve the learning process is modeling and forecasting in the fundamental analysis of the stock market, which indicates that factors in the future of the stock market have an indirect impact. Although the features that have volatile and unstable effects and we have investigated several cases in the research can cause the modeling to be uninterpretable, but the recurrent neural network with LSTM architecture makes it interpretable by modifying the memory. Relying on this issue, we can understand that LSTM interprets the system’s surprising well by correcting the memory. We hope that in the next articles, we will be able to present the method of interpreting the surprising system in LSTM.
Acknowledgments:

We would like to thank M. H. Zhoolideh Haghighi for useful comments and discussion.

[1] Arthur Charpentier Romuald Elie, Carl Remlinger, “Reinforcement Learning in Economics and Finance” 23 April 2021
[2] Celia Escamilla-Rivera, Maryi Alejandra Carvajal Quintero, S. Capozziello “A deep learning approach to cosmological dark energy models” 7 Oct 2019
[3] Wen Guan, Gabriel Perdue, Arthur Pesah, Maria Schuld, Koji Terashi, Sofia Vallecorsa, Jean-Roch Vlimant, “Quantum Machine Learning in High Energy Physics” 18 May 2020
[4] Harry A. Pierson, Michael S. Gashler, “Deep Learning in Robotics: A Review of Recent Research” 22 Jul 2017
[5] Alvarez-Ramírez J, Alvarez J, Rodriguez E. Short-term predictability of crude oil markets: a detrended fluctuation analysis approach. Energy Econ 2008;30: 2645-56.
[6] S. Moshiri, and F. Foroutan, “Forecasting nonlinear crude oil futures prices,” The Energy Journal vol. 27, pp. 81-95, 2005.
[7] Siddhi Vinayak Kulkarni and Imad Haııdar, 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.
[8] Ye, M., Zyren, J., Blumberg, C. and Shore, J., 2008. A Short-Run Crude Oil Price Forecast Model with Ratchet Effect. Atlantic Economic Journal, 37(1), pp.37-50.
[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] 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.
[11] Ling Qi, Matloob Khushi, Josiah Poon,” Event-Driven LSTM For Forex Price Prediction” 29 Jan 2021
[12] Varun Gupta, & Ankit Pandey. (2018). Crude Oil Price Prediction Using LSTM Networks. International Journal of Information, Control and Computer Sciences, 11.0(3).
[13] Cen, Zhongpei, and Jun Wang. ”Crude oil price prediction model with long short term memory deep learning based on prior knowledge data transfer.” Energy 169 (2019): 160-171.
[14] Altan, Aytaç, and Seçkin Karasu. ”Crude oil time series prediction model based on LSTM network with chaotic Henry gas solubility optimization.” Energy (2021): 122964.
[15] Sen, Abhibasu. ”An Analysis of Crude Oil Prices in the Last Decade (2011-2020): With Deep Learning Approach.” Available at SSRN 3906517 (2021).
[16] Arfaoui, Mongi, and Aymen Ben Rejeb. "Oil, gold, US dollar and stock market interde-
pencies: a global analytical insight.” European Journal of Management and Business Economics (2017).
[17] X Chen, S Liu, R Sun, M Hong .”On the convergence of a class of adam-type algorithms for non-convex optimization” arXiv preprint arXiv:1808.02941, 2018.