Forecasting Crude Oil Price with Hybrid Approaches

Amina Ladhari Zahouani[1] and Heni Boubaker[1,2,*]

[1]*Institute of High Commercial Studies (IHEC) of Sousse, LaREMFiQ, B.P. 40, 4054 Sousse, Tunisia.*

[2]*IPAG Business School, IPAG LAB, 184 Boulevard Saint-Germain, 75006 Paris, France.*

Abstract: The objective of this paper is to study the oil price forecasting using parsimonious models that provide the best possible match using the fewest variables or parameters and to compare the performance of our hybrid forecasting models, namely Artificial Neural Network (ANN)-Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN)-LSTM, Bidirectional Recurrent Neural Network (BRNN)-LSTM, and LSTM-Attention, in predicting the daily oil price. This study uses time series data of crude oil price. The data was preprocessed and divided into training and testing sets. The four hybrid models were developed and trained on the training set. The performance of the models is evaluated by using three metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The results show that the LSTM-Attention model outperforms the other models in terms of all three-evaluation metrics, with the lowest error rate. Nevertheless, this approach produces smaller predicting errors than the other computing techniques.

Keywords: Forecast performance, Neural-Networks, Hybrid models, LSTM-Attention, Oil price.

1. INTRODUCTION

Crude oil is a critical commodity in the global economy, with the International Energy Agency (IEA) stating that it makes up more than 31 percent of the world's primary energy and that it is primarily used in transportation, generating electricity and heating. Being a vital component of industry, crude oil is often referred to as the lifeblood of the economy. It plays a significant role in the world economic market. Being an essential energy generator, it is used for a wide range of purposes such as transportation. It is refined to produce gasoline, diesel, and aviation fuel that powers cars, trucks, ships, buses, airplanes and other vehicles. It is also used in industry for the production of plastics, chemicals and cosmetics. It is used even in the medical sector to make antibiotics and hormone treatments. Thus, the price of oil is a key element in the world of economy as it affects the cost of different services around the world. The price of oil can fluctuate rapidly due to a range of factors including changes in global demand, political instability, natural disasters and the actions of major oil producers. Forecasting the daily crude oil price is difficult to undertake due to the volatile nature of the oil market and the varied factors that can affect the price of oil.

Being significantly important, the price of crude oil with its fluctuation plays a major role in the world-wide economy impacting political activities around the world. For this reason, the oil price volatility has always been an important issue to discuss and a major topic for many studies and analyses carried out by economists, analysts and policymakers around the world. It has been a crucial topic for research and understanding. As a result of its fluctuation, forecasting the oil price is a hard and challenging task to be done because it is influenced by a wide range of factors such as the oil demand and supply, the geopolitical tensions and the macro economy. The natural disasters and the political instability also come into play. In fact, the production and supply of oil are controlled by some major countries including Russia, USA, Saudi Arabia and Iran. These nations play a major role in affecting the international price of oil via their production decisions and governmental policies.

Throughout the years, the oil market has witnessed a huge volatility and unpredictability due to various factors that influenced the price and the worldwide economy. In addition, the oil market has faced several crises across the timeline of history. The most remarkable one is the 1973 oil crisis which compelled the members of OAPEC to impose an oil embargo on the nations who supported Israel. The thing which engendered a severe oil shortage impacting the global economy by the surge in the prices of oil.

The second important crisis occurred in 1979. It was caused by the Iranian revolution, the Iran-Iraq conflict and The Soviet Afghan War. Later, the world witnessed a third crisis known as the Gulf War. This began in 1990 when Iraq invaded Kuwait which caused a geopolitical tension and a lot of environmental damage. Indeed, the gulf war initiated an extended period of insecurity and instability in the Middle East. Then, in 2014, there was an oversupply of oil on the worldwide market and the international economy was moving slowly which caused a steep decrease in oil prices.

*Address correspondence to this author at the Institute of High Commercial Studies (IHEC) of Sousse, LaREMFiQ, B.P. 40, 4054 Sousse, Tunisia; E-mail: heniboubaker@gmail.com

Added to all those crises, we can not forget to mention the Corona virus crisis in 2019. And since then, the pandemic has continued to affect our daily life. COVID-19 caused so much damage touching different sectors such as the energy sector and impacted the oil industry. Initially, it produced a significant downturn in demand due to the decrease in global oil demand. Secondly, it resulted in a supply shock due to an oil trade war between the oil-producing countries.

All those events had heavy economic and geopolitical consequences and resulted in extremely volatile oil prices. Indeed, fluctuations in the price of crude oil can be extreme and sudden caused by many factors including demand and supply rates, natural disasters, geopolitical conflicts and changes in government policies. Many studies in the literature make reference to Hamilton's work on oil price volatility. Oil price volatility is a well-known phenomenon in the energy market. These changes in oil price depend on the interaction between oil demand and oil supply, which are as well directly or indirectly impacted by economic conjecture and geopolitical tension (Hamilton (1983)). Therefore, the price of crude oil can be extremely uncertain which makes it difficult to predict and manage risk in the oil market due to this challenging nature.

During the last two decades, the worldwide petroleum market has encountered substantial instability due to the fluctuations in oil prices and many other factors. In the early 2000s, oil prices were quite stable but they began to increase significantly in 2004 as a result of high demand from big nations like India and China. By the year 2008, prices had reached a peak of over 100 dollars per barrel. Then, by the end of the year, the prices witnessed a decrease because of the global financial crisis.

However, between 2009 and 2014, oil prices regained some ground in the years after the financial crisis. But they remained volatile. In late 2014, prices began to decrease owing to many factors such as a spike in oil production in the United States and a demand growth that didn't meet the expectations. The prices experienced a decline of over 100 dollars per barrel in mid-2014 to around 30 dollars per barrel in early 2016. In the beginning of 2016, in an attempt to boost oil prices, OPEC and other important oil-producing nations cooperated to reduce oil production levels. Their measures proved fruitful and the oil prices climbed up to 80 dollars per barrel.

Nonetheless, in 2019, due to the Corona virus crisis, a global economic downturn and a marked decrease in the demand for oil made oil prices fall sharply in the beginning of 2020. According to Albulescu (2020), there is evidence that remarks concerning the wide spread of Corona virus played a role in the fluctuation of oil prices. Oil prices might have recovered in 2021 but supply and demand imbalances remained noticeable in the market along with persistent geopolitical tensions and uncertainties. In general, the crude oil market has been defined by major fluctuations and unpredictability for the past 20 years. That's why forecasting the daily oil price is a tough task to accomplish due to this volatility.

Overall, the importance of oil price forecasts stems from the fact that they are essential not only for stake-holders, such as oil-intensive industries, investors, financial corporations and risk managers, but also for regulators and central banks, in order to measure financial and economic stability (Elder and Serletis (2010)). Actually, oil price prediction has become a critical and severe issue in the area of forecasting research despite the various approaches and techniques that can be used to predict daily oil price. Each possibility has its pros and cons. For example, one known approach which uses the parsimonious model for forecasting oil prices is the Auto Regressive Integrated Moving Average (ARIMA) model. Another approach is to use machine learning algorithms to estimate the oil price such as Random Forests, Neural Networks and Support Vector Machine. There are other algorithms used like Artificial Neural Networks (ANNs), Support Vector Regression (SVR), Random Forests (RF), Gradient Boosting (GB), Random Walk (RW), Error Correction Models (ECM). Additionally, other approaches can be mentioned as Fundamental Analysis or Sentiment Analysis. There is also the Hybrid Model for forecasting oil prices. This integrates two or more different approaches to enhance future oil prices. The purpose of the Hybrid Model is to capitalize the strengths of each model and offset the weaknesses in order to produce more reliable and precise predictions. There is a range of possible combinations including Combine Traditional Statistical and Econometric Techniques with machine learning techniques or Combine Wavelet Transform and Support Vector machines. Using the Hybrid Model in forecasting crude oil prices improves the reliability of the predictions so that investors and analysts make more knowledgeable choices about when to invest in or divest from oil future or stocks.

The following sections of this research paper are organized in the subsequent manner. Section 2 reviews the literature. Section 3 explains some mathematical methods. Section 4 and 5 are for data analysis. Section 6 is for analyzing final results. Finally, section 7 is for concluding the paper.

2. REVIEW OF THE LITERATURE

The significant impact of oil price forecasting on various economic sectors and markets significantly drew the attention of scholars and policymakers. That's why over the past few years, several papers on predicting oil prices have been released. One of the first investigations in this field of research was performed by Professor Jack Knetsch in (2007) in which he used a statistical method of forecasting oil prices. His method has shown a combination of two approaches: the time-series models and statistical modeling. The Knetsch method uses the advantages of each approach to make a closer prediction. Moreover, the author found that the efficient forecasting models should be done in limited time (from one to eleven months).

Besides the time-series analysis and econometric modeling performed by Francesco Coppola (2008), he also uses the Vector Error Correction Models (VECM) as a forecasting approach. VECM is a type of time-series model based on collected data from the future markets. This method is used particularly for the mid-range and long-range forecasts. Withal, Liao and Wong (2010) focused on a kind of machine learning algorithm for oil forecasting research called Neural Network model. In their study they suggested extracting daily crude oil price from the two major global markets of oil: The Brent and WTI from the late 1990s to the late 2000s. According to them, this approach is useful especially during periods of extreme fluctuations. Moreover, Alquist and Kilian (2010) utilized a structural VAR model to examine the predictive value of crude oil future prices for futures spot prices. In order to confirm the reliability of their forecasts, they compared them to other approaches such as the Random Walk and Survey-based models. Based on their results, it appeared that the forecasts derived from crude oil futures prices are less accurate than those generated by the Random Walk.

Furthermore, Arouri et al. (2011) oil price prediction method is Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. It is a statistical method to model and forecast the volatility of several time series. The study focuses on daily crude oil in North Africa and the Middle East between 2000 and 2010 and analyzes the impact of crude oil price changes on stock market volatility. Another study written by Baumeister et al. (2013) suggested a hybrid model for forecasting crude oil prices using both technical and fundamental analysis, in which the author investigated and found that the futures spreads can offer essential predictive analytics of the spot crude oil prices as confirmed by several strength checks. Didier and Inekwe (2013) used as a methodology the ARIMA model to predict Brent crude oil prices. The authors exchanged daily Brent oil price in the period extending from 2000 to 2010 and found out that the ARIMA approach was able to discover the inherent variations.

Additionally, to the combination of the ARIMA and the ANN models to forecast oil prices, Basher and Sadorsky (2016) suggested a technique involving wavelet decomposition and Multiple Linear Regression (MLR) to estimate crude oil prices. The writers used wavelet decomposition on the everyday West Texas Intermediate (WTI) crude oil price data between 1990 and 2014 to identify and separate the various frequency components of the time series. A slightly different approach was adopted by Wang et al. (2021) in which the author put forward a model that combines multiple scales and machine learning to forecast WTI crude oil prices. The approach utilized various machine learning techniques, such as wavelet decomposition to identify the feature of WTI crude oil prices at different scales, Principal Component Analysis (PCA) to simplify the features, and Support Vector Regression (SVR) to develop a forecast model that is based on the simplified features.

In addition to combining the previous models, there are many other hybrid models that can be created by combining multiple deep learning models. For example, some researchers have combined Long Short-Term Memory (LSTM) with other types of Recurrent Neural Networks (RNNs) like Gated Recurrent Units (GRUs) to form a hybrid model. LSTM models can be combined with other deep learning models, such as Convolutional Neural Networks (CNNs) and auto encoders, to form hybrid models that can capture both spatial and temporal patterns in time series data (Zhang et al. (2019)). Still, A deep learning ensemble approach for crude oil price forecasting was proposed by Zhao (2017) named SDAE-B. The method was a combination of several deep learning to boost the precision of crude oil price estimation. In order to increase the validity of his study, the author compared it to varied models such as ARIMA and SVR.

Withal, Kang and Ratti (2018) suggested a Structural VAR (Vector Autoregression) model as a methodology which involves incorporating disaggregated world non-US and US oil supply variables into the Structural VAR model for the oil market developed by Kilian and Park (2009). The writers utilized monthly observations on oil from 1973 to 2014 to determine the Structural VAR model's parameters. In fact, VAR models are widely used for forecasting oil prices. These models rely on vector auto-regressions, which are statistical tools used to model the behavior of multiple variables over time. According to Baumeister and Kilian (2012), VAR-based projections that incorporate oil market fundamentals in a recursive manner produce lower forecast errors than forecasts based on futures and time-series models. As a matter of fact, the Oil price forecasting incorporates both VAR-based and time-series models. However, each model has a different methodology and advantage. To predict future prices, time-series models like ARIMA or GARCH rely on past time-series data. It is common to use them for modeling the long-term trends and fluctuation of oil prices. In contrast, VAR models permit the examination of different variables which can assist in identifying sudden market changes due to their sensitivity to recent data changes. These models are often employed for short-term forecasting.

Recently, Xu et al. (2019) use a Deep Neural Network (DNN) model for estimating crude oil prices. The historical crude oil price data was used to train the model, which was then assessed using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to establish the exactness of the findings. After a range of tests, the study proved that using DNN model has a better performance than the traditional machine learning models in predicting oil prices. More newly, Nguyen et al. (2020) suggested a new approach named Long Short-Term Memory (LSTM) neural network model for forecasting crude oil prices. The authors employed daily data about oil prices and other economic benchmarks to test the model. To validate the accuracy of the findings, they analyzed the effectiveness of the LSTM model in comparison to ARIMA and SVR (Support Vector Regression) and proved that LSTM model delivered more accurate predictions than the traditional models. Once more, Wang et al. (2021) introduced an ensemble learning approach with multiple scales to estimate WTI crude oil prices. The model employs multiple machine learning methods like SVR, RF and XG Boost. These models are used for both short-term and long-term forecasting.

All the studies mentioned above mainly utilize monthly data. However, Baumeister et al. (2015) in their research employed diverse machine learning methods to predict weekly crude oil prices. After collecting the data, they used machine learning models such as Neural Networks, Support Vector Machines, and Random Forests in order to estimate future oil prices.

To sum up, the level of accuracy in forecasting the crude oil prices is influenced by a range of factors like the quantity and the quality of the available data, oil demand and supply as well as the level of accuracy required short-term, mid-term or long-term and many other factors. Based on our needs, and after analyzing available data, we can choose the appropriate method to forecast oil price. After choosing the convenient methods, a lot of tests should be done to evaluate their effectiveness and performance by employing suitable measures such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Scaled Error (MASE) and other standard measures used for statistical evaluation.

It is widely acknowledged in the prediction literature that no single model is optimal for all scenarios. Indeed, there is no model that can predict oil prices with complete accuracy. Hence, it is recommended to use hybrid models for a better prediction. A combination of two or more models may provide better results. To align with the existing literature, our attention in this paper is directed to this latter. In Section 3 we will focus on four different hybrid models mostly used in the last five years. We will try to compare them and choose the best possible combination in order to have the most precise model in its predictions with the lowest error rate.

3. HYBRID FORECASTING MODEL

3.1. Artificial Neural Network (ANN)-(LSTM)

According to Hochreiter and Schmidhuber (1997) the combination of Artificial Neural Networks (ANNs) and Long Short-Term Memory (LSTM) is a powerful approach to solving problems that involve sequential data with long-term dependencies. The integration of Artificial Neural Networks (ANNs) and LSTM is especially advantageous in fields such as speech recognition, natural language processing, and time-series analysis. For instance, in speech recognition, ANNs can be utilized to pre-process the audio signal and extract crucial features, while the LSTM can be employed to model the sequence of phonemes or words, thus allowing the model to understand the spoken language better. In addition, ANN-LSTM model can also be used in financial Forecasting (Guo et al. (2019), Ding et al. (2019)) in order to predict stock prices, exchange rates and commodity prices. Several studies have shown that the hybrid model is better than individual LSTM and ANN models for time series forecasting (Wei et al. (2018) and Shahbazian and Savoji (2020)).

3.1.1. Long Short-Term Memory (LSTM)

LSTM stands for Long Short-Term Memory. It is a type of Recurrent Neural Network (RNN). It can spontaneously save and extract temporal state information. LSTM can be used to model sequential data. According to research studies, LSTM models have found applications in diverse sectors, for example according to Fischer and Krauss (2018) in their study they present the application of LSTM models in finance to predict stock returns and directional movements in financial time series.

The equations for an LSTM model are given by

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \qquad (1)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \qquad (2)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \qquad (3)$$

$$g_t = tanh\left(W_g \cdot [h_{t-1}, x_t] + b_g\right) \qquad (4)$$

$$c_t = f_t c_{t-1} + i_t g_t \qquad (5)$$

$$h_t = o_t \cdot tanh(c_t) \qquad (6)$$

where x_t is the input at time t, h_t is the hidden state at time t, c_t is the cell state at time t, σ is the sigmoid function, tanh is the hyperbolic tangent function.

3.1.2. Artificial Neural Network (ANN)

Artificial Neural Networks (ANNs) are a subset of Artificial Intelligence that have enhanced the way in which machines can replicate human cognitive processes. According to the findings of Naik et al. (2019). The Artificial Neural Network (ANN) is widely recognized as an efficacious AI computing device that has found wide-ranging applications in several fields, such as telecommunications, materials research (Li and Wang (2019), healthcare, neuroscience (Singh et al. (2016)) and finance (Zhang and Patuwo (1998)). ANN is an algorithm used for classification and regression tasks. The ANN technique relies primarily on the input layer and the data-gathering capabilities of the hidden layers to produce the output layer.

The ANN model can be represented as

$$y(t) = f(W_2 f(W_1 x(t) + b_1) + b_2) \qquad (7)$$

where:

- x(t) is the input vector at time t.
- W1 and W2 are the weight matrices for the input-to-hidden and hidden-to-output layers, respectively.
- b1 and b2 are the bias terms for the hidden and output layers, respectively.
- f(.) is the activation function used in the neurons of the ANN.

3.2. Convolutional Neural Networks (CNNs)-(LSTM)

A CNN-LSTM model is a type of neural network architecture that combines Convolutional Neural Networks (CNNs) and LSTM networks. By combining the strengths of Convolutional Neural Networks and Long Short-Term Memory networks, a CNN-LSTM model can effectively capture both spatial and temporal dependencies in the input data. The CNN component identifies important features in the spatial dimension, while the LSTM component captures patterns and dependencies over time, resulting in a powerful sequence processing architecture. CNN-LSTM models have found applications including: computer vision, Speech recognition (Shahroudy et al. (2016)) and Timeseries prediction for forecasting stock prices (Tsantekidis et al. (2017)).

To obtain a comprehensive explanation of CNN-LSTM models, we suggest that the reader refers to research articles and papers that delve into the subject in great detail, such as Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting by Shi et al. (2015) or Convolutional Recurrent Neural Networks for Music Classification through Lee et al. (2017).

A Convolutional Neural Network is a deep learning algorithm that is frequently employed in image analysis and computer vision applications. Its key feature is the ability to extract relevant features from raw input data, such as images, by utilizing convolutional layers (Kim and Kim (2019)). CNNs are a valuable tool for time series prediction as they can exploit local features and weight-sharing to significantly decrease the model's parameter count and enhance its learning efficiency (Qin et al. (2019)). A CNN can be expressed in the following mathematical expression

$$output = f(W \cdot x + b) \quad (8)$$

where:

• $output$ is the output of the neural network,

• f is the activation function used to introduce nonlinearity in the model,

• W is the weight matrix that connects the input to the output,

• x is the input data,

• denotes the matrix multiplication operation,

• b is the bias vector that is added to the output after the matrix multiplication.

Convolutional Neural Networks have been used in various fields for different purposes. Some of the most common applications of CNNs are Computer Vision, Natural Language Processing such as classification, healthcare (Prasad et al. (2019)) and even finance. Besides, Prasad et al. (2019) used a CNN-based model for predicting stock prices using news sentiment data and Passalis et al. (2018) proposed a CNN-based model for detecting anomalies in high-frequency financial data. With their ability to automatically learn and extract features from raw data, CNNs have become a valuable tool for processing complex data across a wide range of fields, and ongoing research is opening up new applications for these techniques.

3.3. Bidirectional Recurrent Neural Network (BRNN)-(LSTM)

A hybrid model that combines bidirectional RNN with LSTM is a robust technique for handling sequential data. It employs two RNN layers, one processing the input sequence in a forward direction and the other in a backward direction. The output layer, connected to both RNN layers, produces the final prediction. Moreover, Graves, A. (2013), by combining the forward and backward LSTM representations at each time step; we obtain a sequence embedding that incorporates past and future contexts. This approach has shown excellent performance on a range of sequence prediction tasks including speech recognition and natural language generation. The inclusion of an LSTM layer between the two RNN layers allows the model to selectively retain or discard information from past time steps and adjust the current state based on the input at the present time step. This feature enables the model to handle long-term dependencies and overcome the issue of vanishing gradients that are commonly found in traditional RNNs. The BRNN-LSTM model has found diverse applications in different domains, including finance. Specifically in finance, the model has been utilized for a range of purposes, such as forecasting stock prices, managing portfolios, detecting fraudulent activities, evaluating credit risks and other relevant areas (Chen et al. (2017) and Fischer et al. (2018)). For example, the BRNN-LSTM model has been employed to forecast stock prices by leveraging past price data and other contextual data such as news articles, company financial reports, and social media sentiments. This approach enables the model to capture the time-varying patterns in stock prices and update its predictions based on new information in real-time (Ntakaris et al. (2020)). The expression for the BRNN-LSTM model can be broken down into three main parts: the forward RNN layer, the backward RNN layer and the LSTM layer.

A Bidirectional Recurrent Neural Network is a neural network architecture that can process input sequences in both forward and backward directions. This allows the network to capture information from the past and the future providing context for each time step in the sequence. A BRNN is composed of two recurrent layers: one processes the input sequence in the forward direction, and the other processes it in the backward direction.

The equations for the forward RNN layer can be expressed as

$$h_{f,1} = f\left(W_f x_1 + b_f\right) \quad (9)$$

$$h_{f,i} = f\left(W_f x_i + U_f h_{f,i-1} + b_f\right), i = 2, \dots, n \quad (10)$$

where f is the activation function (e.g., tanh or ReLU), W_f is the weight matrix for the input, U_f is the weight matrix for the hidden state, and b_f is the bias term. $h_{f,i}$ represents the forward hidden state at time step i and x_i is the input at time step i.

The equations for the backward RNN layer are as follows

$$h_{bn} = f(W_b x_n + b_b) \quad (11)$$

$$h_{bi} = f(W_b x_i + U_b h_{i+1} + b_b), fori = n - 11 \quad (12)$$

where W_b, U_b, and b_b are the weight matrix and bias term for the backward layer, and f is the activation function.

Conforming to Shi et al. (2020), BRNN has been widely used in various domains; including natural language processing such as sentiment analysis, speech recognition, image processing, and even in finance BRNN has been used for applications such as stock price prediction, fraud detection and credit scoring.

3.4. LSTM-Attention

LSTM-attention is a neural network architecture that integrates the Long Short-Term Memory layer and an attention mechanism. It utilizes the LSTM layer to model sequential data and capture long-term dependencies, while the attention mechanism enables the model to emphasize specific parts of the input sequence that are significant for the task. Attention is widely regarded as a critical concept in deep learning. The attention mechanism has found extensive use in various application domains and has been shown to significantly enhance the effectiveness and precision of perceptual information processing. Bahdanau et al. (2014) argues that the attention mechanism enables the model to adjust the relative significance of various input components, assigning more importance to the ones that are more crucial for the given task. This leads to improved performance and enhanced

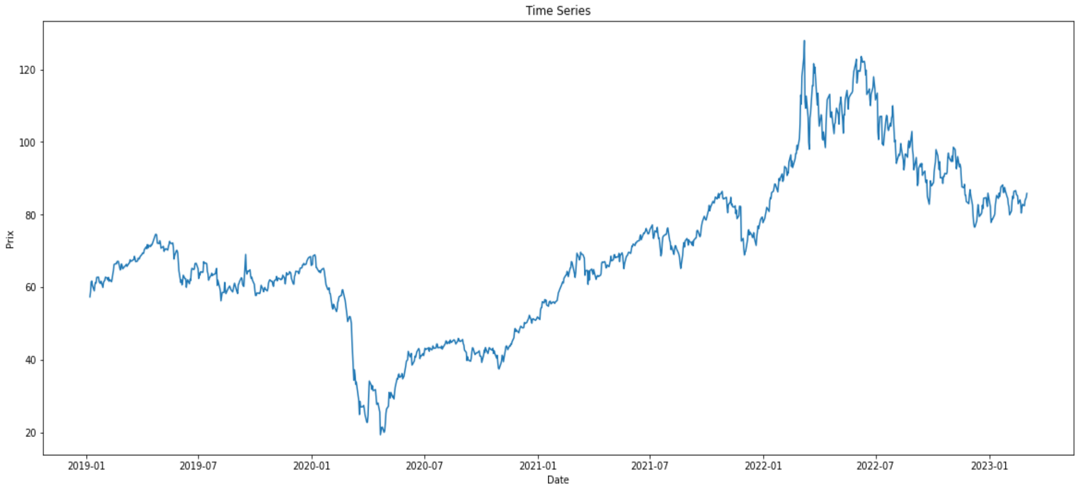

Fig. (1). Time Series Plot of the Monthly Crude Oil Price.

comprehensibility of the model. Additive attention and multiplicative attention are the two most widely adopted mechanisms, with the latter being computationally less demanding. (Luong and Pham (2015)).

4. METHODOLOGY

In this section, we introduce the information utilized in this research. Then, we outline the standards used to evaluate the model in this investigation.

4.1. Dataset

The information utilized in this analysis is centered on the daily price of crude oil (petroleum) and is denominated in dollar per barrel. The time series data for the daily Crude Oil Price was used during the period from 1 July 2019 to 3 March 2023. The dataset contained almost 1000 observations. The first 980 observations were used as the training series, while the last 20 observations were used as the testing series. Fig. (**1**) illustrates the fluctuations in the price of crude oil over time, using a monthly time scale on the x-axis and the corresponding values on the y-axis. This visual representation enables us to discern patterns, identify trends, and interpret the data's behavior.

4.2. Model Evaluation

The process of evaluating a model is just as critical as the development process. Based on the accuracy performance results, it is important to reevaluate the model development process including the selection of a suitable method until the most appropriate model is achieved. In this study, we employed two commonly used error metrics to evaluate the performance of the models.

• Mean Squared Error (MSE) is defined as follows

$$MSE = \frac{1}{n}\sum_{i=1}^{n} \quad (13)$$

• Mean Absolute Error (MAE) is given by

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (14)$$

• Mean Absolute Percentage Error (MAPE) is expressed as follows

$$MAPE = \frac{100}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (15)$$

where n is the number of test data; Y_1 is the actual value and $\hat{Y}_i$ is the predicted value.

5. DETERMINATION PARAMETERS OF MODEL

5.1. Fitting the hybrid CNN-LSTM

The section is dedicated to utilizing the hybrid CNN-LSTM model, which was explained in Section 3, to predict the daily price of Crude Oil. Table **1** displays the predicted and actual values for last 20 observations.

Table 1. Observed and Forecasted Values Using the Hybrid Model CNN-LSTM.

Day	Actual	Forcasted
2023-03-04	42.430000	39.558647
2023-03-05	41.029999	39.659187
2023-03-06	40.950001	39.701328
2023-03-07	40.930000	39.698246
2023-03-08	39.270000	39.659821
2023-03-09	39.943333	39.526436
2023-03-10	40.616665	39.356026
2023-03-11	41.290001	39.181515
2023-03-12	42.650002	39.042973
2023-03-13	41.990002	38.986591
2023-03-14	43.340000	41.258984
2023-03-15	42.849998	41.369743
2023-03-16	42.473331	41.764885

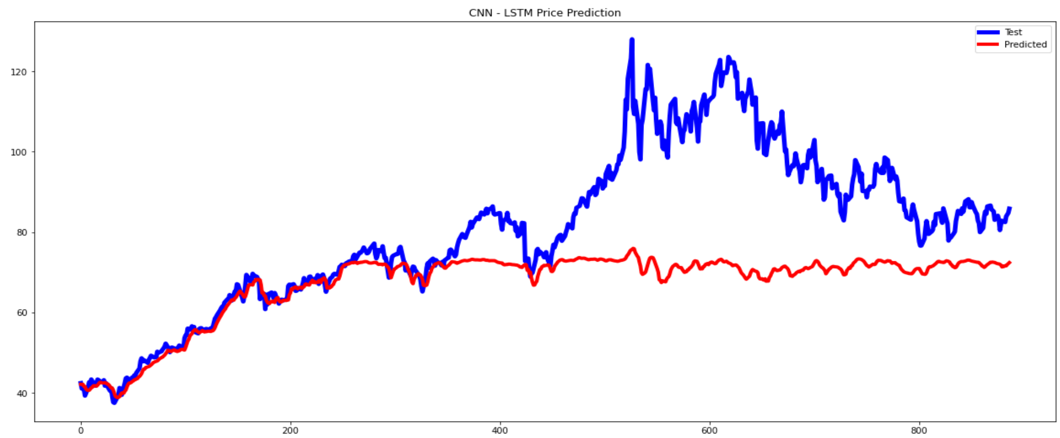

Fig. (2). CNN-LSTM Price Prediction.

2023-03-17	42.096668	41.789139
2023-03-18	41.720001	41.708721
2023-03-19	42.450001	41.640362
2023-03-20	43.320000	41.896152
2023-03-21	43.160000	42.243332
2023-03-22	42.930000	42.586613
2023-03-23	42.826668	42.611725

Fig. (**2**) represents the daily crude oil price prediction in the period from 1 July 2019 to 2 August 2022, giving a total of 800 observations. As illustrated in Fig. (**2**), both the predicted values of CNN-LSTM model and the actual values are shown, allowing for a direct comparison to evaluate the model's performance.

To assess the ability of the forecasting models, they are utilized for daily Crude Oil price prediction. The performance evaluation metrics used in this study include three measures: Mean Square Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Table **2** exhibits the achieved forecasting accuracy of outcomes in the hybrid model with regard to MSE, MAP and MAPE.

Table 2. The Obtained Prediction Performance Results.

Model	MSE	MAE	MAP	Accuracy
CNN - LSTM	388.8773	13.584178	0.13997169	99.81

5.2. Fitting the Hybrid ANN-LSTM

This segment focuses on using the hybrid ANN-LSTM model, which was previously described, to forecast the daily price of Crude Oil. Table **3** exhibits the anticipated and factual values for the latest data points. Fig. (**4**) depicts the daily crude oil price prediction for the same time period as previously mentioned (1 July 2019 to 2 August 2022). The figure displays both the predicted values of the ANN-LSTM model and the actual values. According to both Figs. (**3** and **4**), the ANN-LSTM model outperforms the CNN-LSTM model in terms of predictive accuracy.

Table 3. Observed and Forecasted Values Using the Hybrid Model ANN-LSTM.

Day	Actual	Forecasted
2023-03-04	42.430000	40.234219
2023-03-05	41.029999	40.327461
2023-03-06	40.950001	40.383930
2023-03-07	40.930000	40.407757
2023-03-08	39.270000	40.403210
2023-03-09	39.943333	40.330048
2023-03-10	40.616665	40.223934
2023-03-11	41.290001	40.102612
2023-03-12	42.650002	39.995605
2023-03-13	41.990002	39.938564
2023-03-14	43.340000	39.910053
2023-03-15	42.849998	39.930210
2023-03-16	42.473331	39.994324
2023-03-17	42.096668	40.076035
2023-03-18	41.720001	40.159420
2023-03-19	42.450001	40.229881
2023-03-20	43.320000	40.303333
2023-03-21	43.160000	40.395691
2023-03-22	42.930000	40.501076
2023-03-23	42.826668	40.614197

Table 4. The Obtained Prediction Performance Results.

Model	MSE	MAE	MAP	Accuracy
ANN - LSTM	699.98645	14.561506	0.14855464	99.64

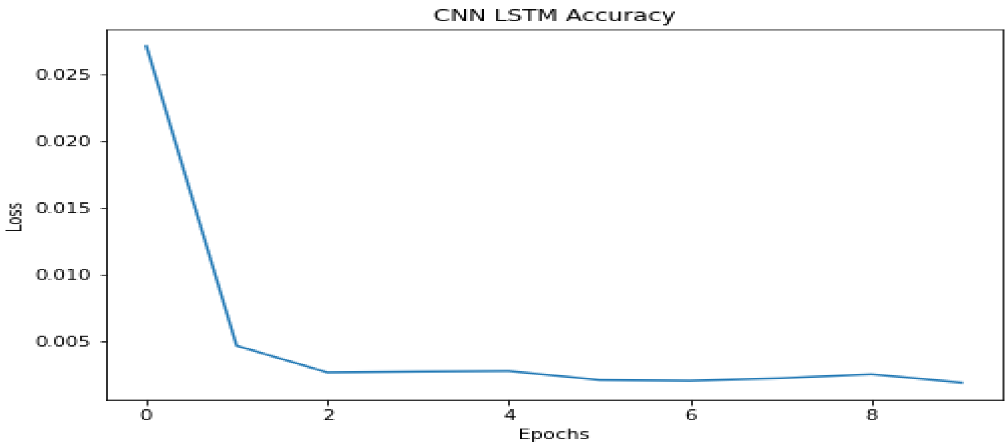

Fig. (3). CNN LSTM Accuracy.

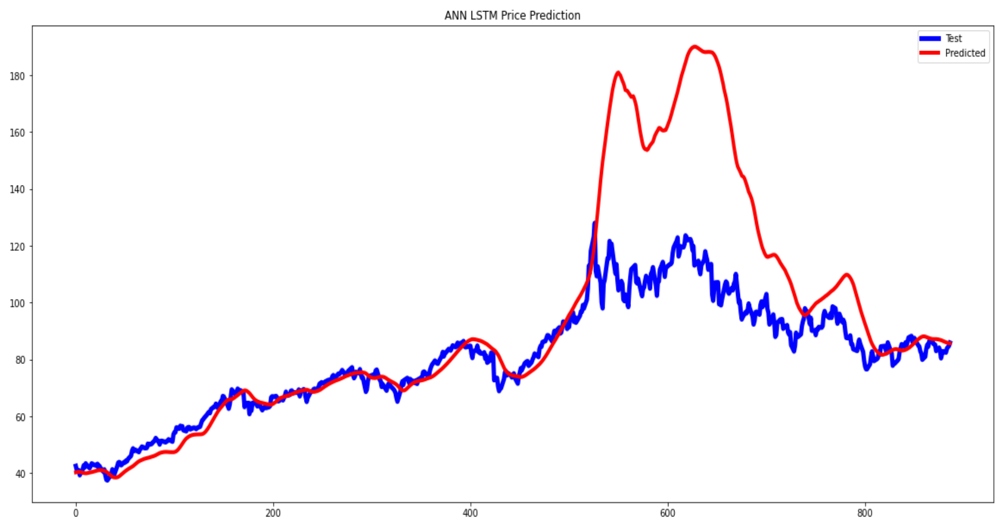

Fig. (4). ANN-LSTM Price Prediction.

5.3. Fitting the hybrid BRNN-LSTM

This section discusses the application of the hybrid BRNN-LSTM model, as described in Section 3, for predicting the daily price of crude oil. Table **5** shows the predicted and actual values for the most recent data points.

Table 5. Observed and Forecasted Values Using the Hybrid Model BRNN-LSTM.

Day	Actual	Forcasted
2023-03-04	42.430000	40.855492
2023-03-05	41.029999	40.976685
2023-03-06	40.950001	40.957672
2023-03-07	40.930000	40.904335
2023-03-08	39.270000	40.830383
2023-03-09	39.943333	40.582901
2023-03-10	40.616665	40.369156
2023-03-11	41.290001	40.202778
2023-03-12	42.650002	40.120449
2023-03-13	41.990002	40.188381
2023-03-14	43.340000	40.230373
2023-03-15	42.849998	40.407223
2023-03-16	42.473331	40.574863
2023-03-17	42.096668	40.711979
2023-03-18	41.720001	40.814648
2023-03-19	42.450001	40.863037
2023-03-20	43.320000	40.968575
2023-03-21	43.160000	41.139061
2023-03-22	42.930000	41.294689
2023-03-23	42.826668	41.429363

Fig. (**6**). shows the daily crude oil price prediction using BRNN-LSTM model over the same frame mentioned earlier, covering a total of 800 observations. Based on Figs. (**2**, **4** and **6**), it appears that BRNN-LSTM model presents superior

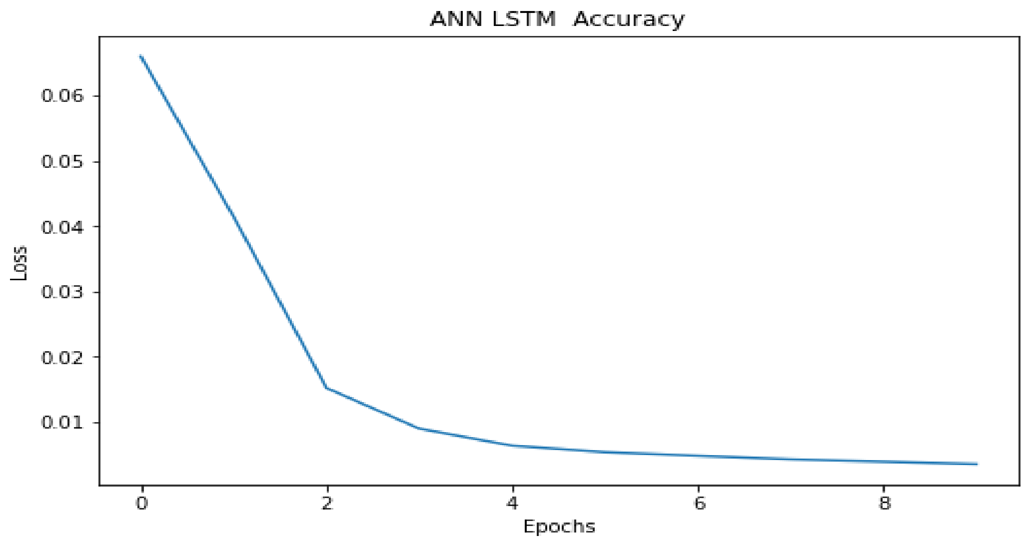

Fig. (5). ANN-LSTM Accuracy.

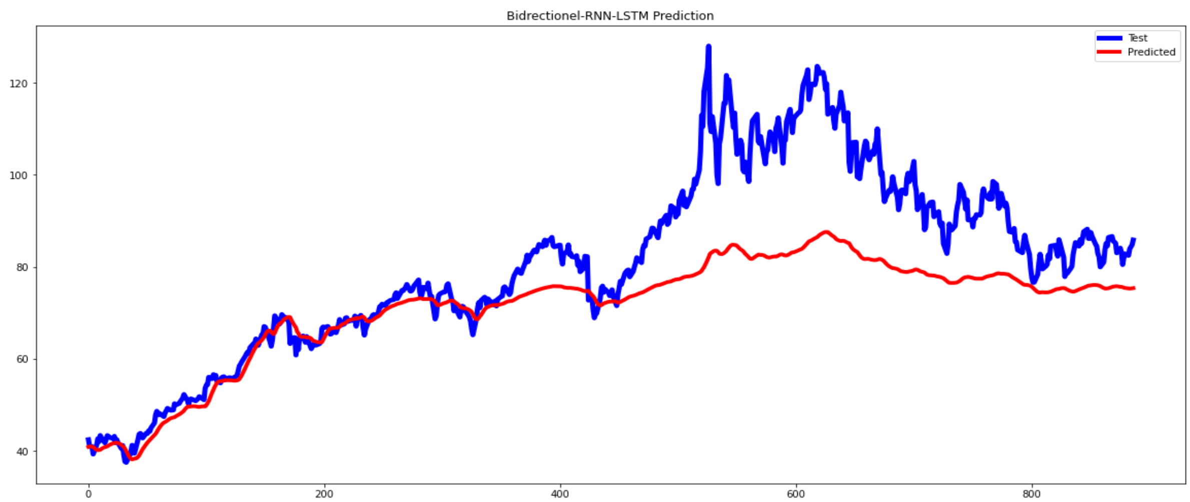

Fig. (6). BRNN-LSTM Price Prediction.

performance in predicting crude oil prices compared to other models.

Table 6. The Obtained Prediction Performance Results.

Model	MSE	MAE	MAP	Accuracy
BRNN-LSTM	193.664	9.798892	0.10306074	99.78

5.4. Fitting the hybrid LSTM-Attention

In this section, we explore the use of the hybrid LSTM-Attention model, which was explained in Section3, to forecast the daily price of crude oil. Table **7** displays the actual and forecasted values for the latest data points.

Table 7. Observed and Forecasted Values Using the Hybrid Model LSTM-Attention.

Day	Actual	Forcasted
2023-03-04	42.430000	41.532707
2023-03-05	41.029999	41.621849
2023-03-06	40.950001	41.677410
2023-03-07	40.930000	41.704960
2023-03-08	39.270000	41.705986
2023-03-09	39.943333	41.646206
2023-03-10	40.616665	41.554832
2023-03-11	41.290001	41.453758
2023-03-12	42.650002	41.364067
2023-03-13	41.990002	41.316906
2023-03-14	43.340000	41.292301
2023-03-15	42.849998	41.309372
2023-03-16	42.473331	41.357086
2023-03-17	42.096668	41.417587

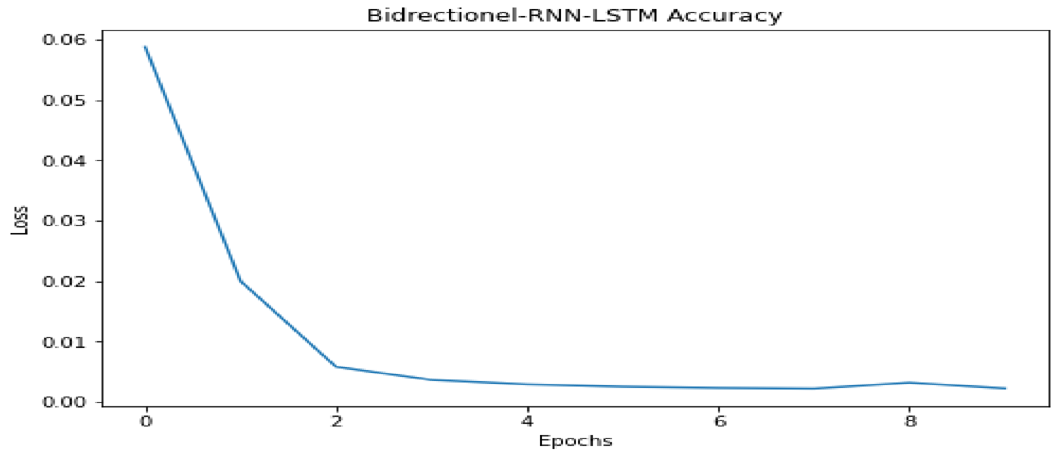

Fig. (7). BRNN-LSTM Accuracy

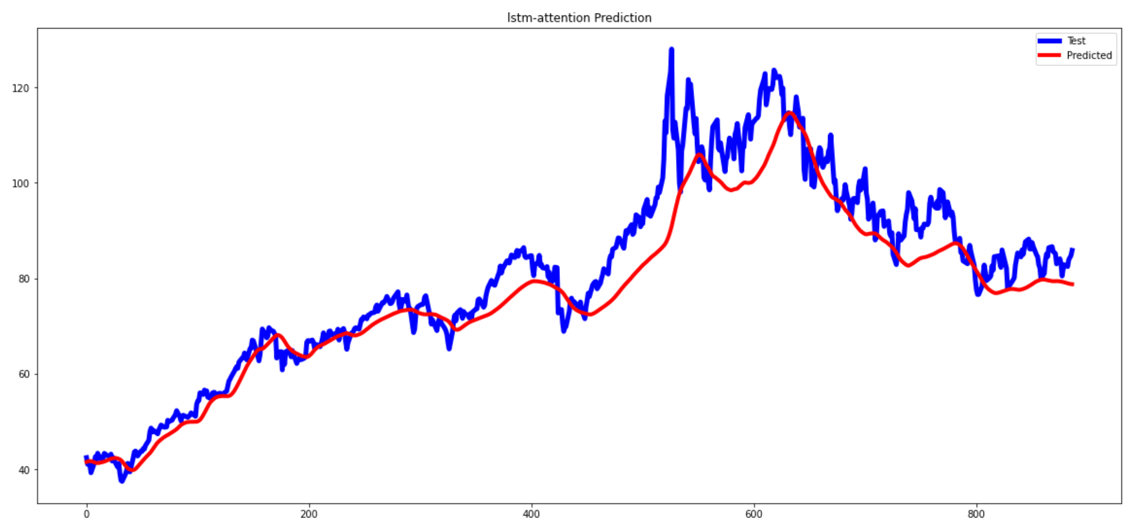

Fig. (8). LSTM-Attention Price Prediction.

2023-03-18	41.720001	41.480743
2023-03-19	42.450001	41.539112
2023-03-20	43.320000	41.605721
2023-03-21	43.160000	41.695675
2023-03-22	42.930000	41.797848
2023-03-23	42.826668	41.904472

Fig. (**8**). depicts the predicted daily crude oil price using LSTM-attention model and the actual crude oil prices from 1 July 2019 to 2 March 2023. The figure shows that LSTM-attention model appears to accurately capture the fluctuations in crude oil prices over time. Additionally, the predicted prices seem to closely follow the actual prices, with only slight deviations in some areas.

Table 8. The Obtained Prediction Performance Results.

Model	MSE	MAE	MAP	Accuracy
LSTM-Attention	45.14683	4.918804	0.056444895	99.55

6. RESULTS

To assess how well the predictive models can predict Crude Oil prices, three prediction performance measures are employed in this study: Mean Square Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Additionally, Figs. (**3**, **5**, **7** and **9**). display the accuracy with the x-axis representing epochs and y-axis representing loss. These figures demonstrate that the model with the lowest loss value tends to have the highest performance.

Table **9** displays the achieved results for prediction performance through CNN-LSTM, ANN-LSTM, BRNN-LSTM, LSTM-Attention model in term of three standards with their accuracy. Table **9** provides clear evidence that the LSTM-Attention model has achieved fewer errors compared to the other models.

Table 9. The Obtained Prediction Performance Results.

Model	MSE	MAE	MAP	Accuracy
CNN - LSTM	388.8773	13.584178	0.13997169	99.81
ANN - LSTM	699.98645	14.561506	0.14855464	99.64

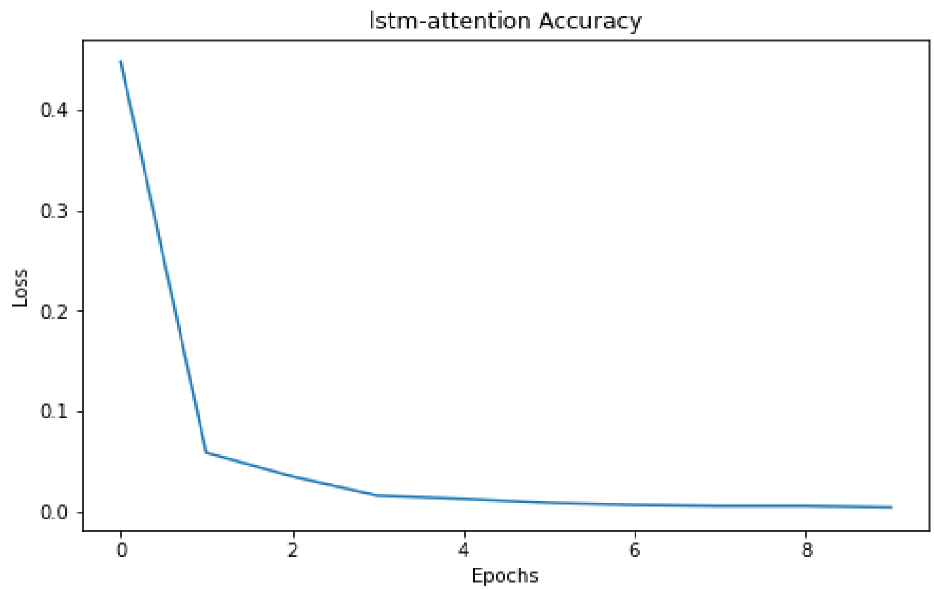

Fig. (9). LSTM-Attention Accuracy.

BRNN-LSTM	193.664	9.798892	0.10306074	99.78
LSTM-Attention	45.14683	4.918804	0.056444895	99.55

The hybrid LSTM-attention model was observed to exhibit better forecasting performance, as it was able to generate predictions that were closer to the actual values and followed a similar pattern as the real data. On the other hand, the other models displayed inferior forecasting performance, particularly when compared to the LSTM-attention model.

7. CONCLUSION

In summary, predicting crude oil prices is a demanding and intricate task to be done due to numerous factors that may impact the market. While conventional methods like Time Series Analysis and Econometric Modeling achieved some success, recent advancements in Machine Learning and Artificial Intelligence have presented novel opportunities to create more precise and trustworthy forecasting models. Within this framework, hybrid approaches that blend various modeling techniques have demonstrated significant potential in enhancing the precision of crude oil price forecasts. Merging statistical models with machine learning algorithms like Artificial Neural networks and Support Vector Regression can capture the nonlinear patterns in the market, yielding, thereby, more sturdy and dependable predictions. The importance of oil prices to the global economy has prompted numerous researchers to develop accurate forecasts of this critical energy commodity. This is because policy-makers aim to be equipped to deal with unexpected price shocks and reduce the potential catastrophic effects on the economy.

In this paper we aim to contrast the CNN-LSTM, ANN-LSTM, BRNN-LSTM and LSTM-Attention models in terms of predicting the Crude Oil (petroleum) daily Price per Barrel. The results obtained from using these models were evaluated by analyzing the (MSE, MAE, and MAPE) outcomes. The process began by collecting time series of daily oil price, and concluded with conducting pair wise comparisons of their predictions' outcomes. The results of our analysis indicate that the hybrid LSTM-Attention model outperforms other hybrid models in terms of accuracy, as evidenced by the lowest error rate. Compared to the CNN-LSTM, ANN-LSTM, and BRNN-LSTM models, the LSTM-Attention model demonstrated a more efficient ability to predict daily Crude Oil (petroleum) price. Through our study, we tried to provide valuable insights into the usefulness of hybrid modeling techniques and the importance of incorporating advanced deep learning models such as the LSTM-Attention in crude oil price forecasting. We firmly believe that the findings can be highly significant for various stakeholders such as investors, traders, and policymakers who rely on accurate forecasting to be able to make informed decisions. Despite the success of our hybrid LSTM-Attention model, there is always room for further improvement. In the future, we will explore the possibility of incorporating additional models and refining our previous hybrid models in order to enhance the accuracy of our forecasts. By doing so, we hope to contribute to the development of more advanced and reliable forecasting models for crude oil prices. We aspire to enhance the precision of our predictions and guarantee the most dependable forecasts that could be attained.

REFERENCES

Lu, Wenjie and Li, Jiazheng and Li, Yifan and Sun, Aijun and Wang, Jingyang. A CNN-LSTM-based model to forecast stock prices. *Complexity*, vol. 2020, pp. 1–10, 2020.

Speiser, Jaime Lynn and Miller, Michael E and Tooze, Janet and Ip, Edward. A comparison of random forest variable selection methods for classification prediction modeling. *Expert systems with applications*, vol. 134, pp. 93–101, 2019.

Guermoui, Mawloud and Melgani, Farid and Gairaa, Kacem and Mekhalfi, Mohamed Lamine. A comprehensive review of hybrid models for solar radiation forecasting. *Journal of Cleaner Production*, vol. 258, pp. 120357, 2020.

Al Mamun, Abdullah and Sohel, Md and Mohammad, Naeem and Sunny, Md Samiul Haque and Dipta, Debopriya Roy and Hossain, Eklas. A comprehensive review of the load forecasting techniques using single and hybrid predictive models. *IEEE Access*, vol. 8, pp. 134911–134939, 2020.

YALÇINER, Ayten YILMAZ and ÖZCAN, Onur. A FORECASTING MODEL PROPOSAL ON RENEWABLE ENERGY PRODUCTION AND FUNDS . *Journal of Life Economics*, vol. 8, pp. 263, 2021.

Belmahdi, Brahim and Louzazni, Mohamed and Bouardi, Abdelmajid El. A hybrid ARIMA–ANN method to forecast daily global solar radiation in three different cities in Morocco. *The European Physical Journal Plus*, vol. 135, pp. 1–23, 2020.

Dudek, Grzegorz and Pełka, Paweł and Smyl, Slawek. A Hybrid Residual Dilated LSTM and Exponential Smoothing Model for Midterm Electric Load Forecasting. *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, pp. 2879–2891, 2021.

Abdollahi, Hooman and Ebrahimi, Seyed Babak. A new hybrid model for forecasting Brent crude oil price. *Energy*, vol. 200, pp. 117520, 2020.

Qian, Zheng and Pei, Yan and Zareipour, Hamidreza and Chen, Niya. A review and discussion of decomposition-based hybrid models for wind energy forecasting applications. *Applied energy*, vol. 235, pp. 939–953, 2019.

Niu, Zhaoyang and Zhong, Guoqiang and Yu, Hui. A review on the attention mechanism of deep learning. *Neurocomputing*, vol. 452, pp. 48–62, 2021.

Xie, Anqi and Yang, Hao and Chen, Jing and Sheng, Li and Zhang, Qian. A Short-Term Wind Speed Forecasting Model Based on a Multi-Variable Long Short-Term Memory Network. *Atmosphere*, vol. 12, pp. 651, 2021.

Haslbeck, Jonas MB and Bringmann, Laura F and Waldorp, Lourens J. A Tutorial on Estimating Time-Varying Vector Autoregressive Models. *Multivariate Behavioral Research*, vol. 56, pp. 120–149, 2021.

Tilk, Ottokar and Alumäe, Tanel. Bidirectional Recurrent Neural Network with Attention Mechanism for Punctuation Restoration. *Interspeech*, pp. 3047–3051, 2016.

Schuster, Mike and Paliwal, Kuldip K. Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, vol. 45, pp. 2673–2681, 1997.

Aldalbahi, Adel and Shahabi, Farzad and Jasim, Mohammed. BRNN-LSTM for Initial Access in Millimeter Wave Communications *Electronics*, vol. 10, pp. 1505, 2021.

Hong, Harrison and Li, Frank Weikai and Xu, Jiangmin. Climate risks and market efficiency. *Journal of econometrics*, vol. 208, pp. 265–281, 2019.

Shi, Heng and Xu, Minghao and Li, Ran. Deep Learning for Household Load Forecasting—A Novel Pooling Deep RNN. *IEEE Transactions on Smart Grid*, vol. 9, pp. 5271–5280, 2017.

Chong, Eunsuk and Han, Chulwoo and Park, Frank C. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, vol. 83, pp. 187–205, 2017.

Fischer, Thomas and Krauss, Christopher. Deep learning with long short-term memory networks for financial market predictions. *European journal of operational research*, vol. 270, pp. 654–669, 2018.

Khan, Muhammad Adnan and Saqib, Shazia and Alyas, Tahir and Rehman, Anees Ur and Saeed, Yousaf and Zeb, Asim and Zareei, Mahdi and Mohamed, Ehab Mahmoud. Effective Demand Forecasting Model Using Business Intelligence Empowered with Machine Learning. *IEEE Access*, vol. 8, pp. 116013–116023, 2020.

Tschora, Léonard and Pierre, Erwan and Plantevit, Marc and Robardet, Céline. Electricity price forecasting on the day-ahead market using machine learning. *Journal of econometrics*, vol. 313, pp. 118752, 2022.

Sezer, Omer Berat and Gudelek, Mehmet Ugur and Ozbayoglu, Ahmet Murat. Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied soft computing*, vol. 90, pp. 106181, 2020.

Liu, Huan and Li, Chenxi and Shao, Yingqi and Zhang, Xin and Zhai, Zhao and Wang, Xing and Qi, Xinye and Wang, Jiahui and Hao, Yanhua and Wu, Qunhong and others. Forecast of the trend in incidence of acute hemorrhagic conjunctivitis in China from 2011–2019 using the Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ETS) models. *Journal of infection and public health*, vol. 13, pp. 287–294, 2020.

Canova, Fabio and Ciccarelli, Matteo. Forecasting and turning point predictions in a Bayesian panel VAR model. *Journal of Econometrics*, vol. 120, pp. 327–359, 2004.

Zhang, Yaojie and Wahab, MIM and Wang, Yudong. Forecasting crude oil market volatility using variable selection and common factor. *International Journal of Forecasting*, vol. 39, pp. 486–502, 2023.

Alrweili, Hleil and Fawzy, Haitham. Forecasting Crude Oil Prices Using an ARIMA-ANN Hybrid Model. *J Stat Appl Probab*, vol. 11, pp. 845–855, 2022.

Kerriche, Nawel and Moussi, Oum El Kheir. Forecasting future natural gas demand in Algeria using Ba Yesian Model Averaging. *les cahiers du cread*, vol. 37, pp. 67–102, 2021.

Zha, Wenshu and Liu, Yuping and Wan, Yujin and Luo, Ruilan and Li, Daolun and Yang, Shan and Xu, Yanmei. Forecasting monthly gas field production based on the CNN-LSTM model. *Energy*124889, pp. 124889, 2022.

Degiannakis, Stavros and Filis, George. Forecasting oil prices: High-frequency financial data are indeed useful. *Energy Economics*, vol. 76, pp. 388–402, 2018.

de Medeiros, Rennan Kertlly and da Nóbrega Besarria, Cássio and de Jesus, Diego Pitta and de Albuquerquemello, Vinicius Phillipe. Forecasting oil prices: New approaches. *Energy*, vol. 238, pp. 121968, 2022.

Mensi, Walid and Sensoy, Ahmet and Vo, Xuan Vinh and Kang, Sang Hoon. Impact of COVID-19 outbreak on asymmetric multifractality of gold and oil prices. *Resources Policy*, vol. 69, pp. 101829, 2020.

Büyükşahin, Ümit Çavuş, Ertekin, Şeyda . Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. *Neurocomputing*, vol. 361, pp. 151–163, 2019.

Li, Xiaodong and Wu, Pangjing and Wang, Wenpeng . Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong. *Information Processing and Management*, vol. 57, pp. 102212, 2020.

Yonar, Harun and Yonar, Aynur and Tekindal, Mustafa Agah and Tekindal, Melike . Modeling and Forecasting for the number of cases of the COVID-19 pandemic with the Curve Estimation Models, the Box-Jenkins and Exponential Smoothing Methods. *EJMO*, vol. 4, pp. 160–165, 2020.

Khan, Firdos and Saeed, Alia and Ali, Shaukat. Modelling and forecasting of new cases, deaths and recover cases of COVID-19 by using Vector Autoregressive model in Pakistan. *Chaos, Solitons & Fractals*, vol. 140, pp. 110189, 2020.

Lütkepohl, Helmut. New Introduction to Multiple Time Series Analysis. *Springer Science & Business Media*, 2005.

Liu, Yanli and Wang, Yourong and Zhang, Jian. New Machine Learning Algorithm: Random Forest. *Information Computing and Applications: Third International Conference, ICICA 2012, Chengde, China, September 14-16, 2012. Proceedings 3*, pp. 246–252, 2012.

Safari, Ali and Davallou, Maryam. Oil price forecasting using a hybrid model. *Energy*, vol. 148, pp. 49–58, 2018.

Ren, Xiaohang and Jin, Chenglu and Lin, Ruya and others. Oil price uncertainty and enterprise total factor productivity: Evidence from China. *International Review of Economics & Finance*, vol. 83, pp. 201–218, 2023.

Bourghelle, David and Jawadi, Fredj and Rozin, Philippe. Oil price volatility in the context of Covid-19. *International Economics*, vol. 167, pp. 39–49, 2021.

Belmahdi, Brahim and Louzazni, Mohamed and El Bouardi, Abdelmajid . One month-ahead forecasting of mean daily global solar radiation using time series models. *Optik*, vol. 219, pp. 165207, 2020.

Dbouk, Wassim and Jamali, Ibrahim. Predicting daily oil prices: Linear and non-linear models. *Research in International Business and Finance*, vol. 46, pp. 149–165, 2018.

Shafie-Khah, M and Moghaddam, M Parsa and Sheikh-El-Eslami, MK. Price forecasting of day-ahead electricity markets using a hybrid forecast method. *Energy Conversion and Management*, vol. 52, pp. 2165–2169, 2011.

Cutler, D Richard and Edwards Jr, Thomas C and Beard, Karen H and Cutler, Adele and Hess, Kyle T and Gibson, Jacob and Lawler, Joshua J. Random forests for classification in ecology. *Ecology*, vol. 88, pp. 2783–2792, 2007.

Yu, Wennian and Kim, II Yong and Mechefske, Chris. Remaining useful life estimation using a bidirectional recurrent neural network based autoencoder scheme. *Mechanical Systems and Signal Processing*, vol. 129, pp. 764–780, 2019.

Zang, Haixiang and Liu, Ling and Sun, Li and Cheng, Lilin and Wei, Zhinong and Sun, Guoqiang. Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations. *Renewable Energy*, vol. 160, pp. 26–41, 2020.

Tang, Xianlun and Dai, Yuyan and Wang, Ting and Chen, Yingjie. Short-term power load forecasting based on multi-layer bidirectional recurrent neural network. *IET Generation, Transmission & Distribution*, vol. 13, pp. 3847–3854, 2019.

Benali, Lamara and Notton, Gilles and Fouilloy, A and Voyant, Cyril and Dizene, Rabah. Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components. *Renewable energy*, vol. 132, pp. 871–884, 2019.

Sun, Yunpeng and Gao, Pengpeng and Raza, Syed Ali and Shah, Nida and Sharif, Arshian. The asymmetric effects of oil price shocks on the world food prices: Fresh evidence from quantile-on-quantile regression approach. *Energy*, vol. 270, pp. 126812, 2023.

Ewing, Bradley T and Kang, Wensheng and Ratti, Ronald A. The dynamic effects of oil supply shocks on the US stock market returns of upstream oil and gas companies. *Energy Economics*, vol. 72, pp. 505–516, 2018.

Sun, Zhao-Yong and Huang, Wei-Chiao. The effects of unexpected crude oil price shocks on Chinese stock markets. *Economic Change and Restructuring*, pp. 1–15, 2023.

Satrio, Christophorus Beneditto Aditya and Darmawan, William and Nadia, Bellatasya Unrica and Hanafiah, Novita. Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET. *Procedia Computer Science*, vol. 179, pp. 524–532, 2021.

Sheridan, Robert P. Using Random Forest to Model the Domain Applicability of Another Random Forest Model. *Journal of chemical information and modeling*, vol. 53, pp. 2837–2850, 2013.

Zhang, Yaojie and Ma, Feng and Wang, Yudong. Forecasting crude oil prices with a large set of predictors: Can LASSO select powerful predictors? *Journal of Empirical Finance*, vol. 54, pp. 97–117, 2019.

Zhang, Yaojie and Wei, Yu and Zhang, Yi and Jin, Daxiang. Forecasting oil price volatility: Forecast combination versus shrinkage method *Energy Economics*, vol. 80, pp. 423–433, 2019.

Lin, Boqiang and Bai, Rui. Oil prices and economic policy uncertainty: Evidence from global, oil importers, and exporters' perspective *Research in International Business and Finance*, vol. 56, pp. 101357, 2021.

Bekiros, Stelios and Gupta, Rangan and Paccagnini, Alessia. Oil price forecastability and economic uncertainty *Economics Letters*, vol. 132, pp. 125–128, 2015.

Lu, Quanying and Li, Yuze and Chai, Jian and Wang, Shouyang. Crude oil price analysis and forecasting: A perspective of new triangle *Energy Economics*, vol. 87, pp. 104721, 2020.

Received: Mar 10, 2023 Revised: Mar 14, 2023 Accepted: May 22, 2023