A Novel Approach to Forecast Crude Oil Prices Using Machine Learning and Technical Indicators

Kshitij A. Kakade [1], Kshitish S. Ghate [1], Raj K. Jaiswal [2,*], and Ritika Jaiswal [3]

[1] Department of Computer Science and Information Systems, Birla Institute of Technology and Science Pilani, KK Birla Goa Campus, India; Email: {kshitijkakade2705, ghatekshitish}@gmail.com (K.A.K., K.S.G.)
[2] National Forensic Sciences University, Goa Campus, India
[3] Department of Economics & Finance, Birla Institute of Technology and Science Pilani, KK Birla Goa Campus, India; Email: ritikaj@goa.bits-pilani.ac.in (R.J.)
*Correspondence: jaiswal.raaj@gmail.com (R.K.J.)

***Abstract*—This study proposes to use a hybrid ensemble learning approach to improve the prediction efficiency of crude oil prices. It combines the Long Short-Term Memory (LSTM) with factors that influence the price of crude oil. The information from fundamental and technical indicators is considered along with statistical model predictions like autoregressive integrated moving average (ARIMA)to make one-step-ahead crude oil price predictions. A Principal Component Analysis (PCA) approach is employed to transform the explanatory variables. This study combines the LSTM with PCA, jointly known as the LP model wherein PCA transforms of the fundamental and technical indicators are used as inputs to improve LSTM predictions. Further, it attempts to improve these predictions by introducing the LSTM+PCA+ARIMA (LPA) model, which uses an ensemble learning approach to utilize the forecast from the ARIMA model, as an additional input. Among LP and LPA models, the LSTM model is used as a benchmark to evaluate the performance of the hybrid models. Based on the result, a significant improvement is seen in the LP model over the chosen window sizes and error metrics. On the other hand, the LPA model performs better across all dimensions with an average improvement of 41% over the LSTM model in terms of forecasting accuracy. Moreover, the equivalence of forecasting accuracy is tested using the Diebold-Mariano and Wilcoxon signed-rank tests.**

***Keywords*—Long Short-Term Memory (LSTM), Principal Component Analysis (PCA), ensemble learning, crude oil, forecasting**

I. INTRODUCTION

Crude oil is an indispensable non-renewable commodity, responsible for meeting nearly a third of the global energy demand [1]. It has far-reaching industrial uses and is one of the most actively traded commodities that exhibit significant volatility in its prices [2]. West Texas Intermediate (WTI), which is traded on the New York Stock Exchange, is widely regarded as the global benchmark of oil trading due to the strength of the USA crude oil buyers, along with the global influence of the New York Exchange [3].

Oil price changes have significant implications for macroeconomic conditions. A substantial rise in oil prices indicates inflation and subsequent recession for countries that import oil, while falling prices may be detrimental to the economic growth of oil-exporting nations. A study by Katircioglu *et al*. showed that oil-price fluctuations negatively impact the GDP, CPI, and unemployment in Organisation for Economic Co-operation and Development (OECD) countries in the long term [4]. Reboredo and Ugolini found a significant impact of the oil market on the stock market for three developed and five BRICS countries [5]. Price fluctuations have been increasing with economic globalization and liberalization, which has added to the overall revenue risk [6]. A combination of supply, demand, inventory, and non-fundamental parameters such as the exchange rate and interest rate make the price prediction of crude oil complex [7, 8].

This study aims to develop a deep learning framework that incorporates fundamental and technical indicators along with predictions from statistical models to forecast one day ahead crude oil prices. The fundamental variables include energy indices, stock prices of major oil companies, interest and exchange rates, price of substitute energy products, and other assets sharing strong relationships with crude oil. Different technical indicators such as the Simple Moving Average Crossover (SMA), Relative Strength Index (RSI), Rate of change (ROC), Moving Average Convergence Divergence (MACD), and Bollinger Band Squeeze are also included in the model.

A PCA transformation is used to remove the impact of multicollinearity and reduce the input dimension of the LSTM network. We introduce the LSTM+PCA (LP) model, which includes the principal components of the transformed fundamental and technical indicators mentioned above as inputs. We also develop the

Manuscript received June 5, 2022; revised July 1, 2022; accepted July 18, 2022; published April 4, 2023.

doi: 10.12720/jait.14.2.302-310

LSTM+PCA+ARIMA (LPA) model by extending the LP model in combination with the ARIMA model's forecasting abilities to enhance the accuracy of model predictions. The ARIMA forecasts are calculated using a 252-day rolling window over the selected period and are combined with the LP model as an additional explanatory variable. An ensemble learning approach that involves combining multiple learning algorithms to obtain their collective performance generates the LPA framework. There are only limited studies that forecast crude oil prices based on information from fundamental, technical, and statistical variables using a deep learning model.

The proposed model efficiency was evaluated over four different window sizes (3, 5, 7, and 11 days). The root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) metrics are used to evaluate the forecasting performance of the models for one-step-ahead crude oil prices. The results of the study strongly confirm our hypotheses. The LP model improves performance compared to the baseline LSTM framework over all chosen window sizes and error metrics. The LPA is the best performing model across all dimensions by a significant margin. We also conduct the Diebold-Mariano (DM) and Wilcoxon signed-rank (WS) tests to compare the predictive accuracy of the two forecasting models.

The remaining section of the paper is organized as follows: Section II discusses the existing literature, while Section III outlines the data used in the study. Section IV discusses the methodology used, and Section V presents the experimental setup. The analysis of the results is done in Section VI, while Section VII concludes the study.

II. Related Work

Predictive frameworks for asset prices in financial studies are generally modelled based on fundamental or technical variables. Dees *et al.* consider supply and demand factors such as OPEC production, production capacity, oil inventories, and demand while modelling the price of crude oil [9]. Similarly, Baumeister and Kilian utilize fundamental economic variables to develop six models to forecast commodity prices accurately [10]. Miao *et al.* tests the significance of six different categories of variables for oil price forecasting using a LASSO model with various supply, demand, financial, and commodity market variables [7].

Technical analysis estimates the direction of asset prices from trading activity, such as price movement and volume [11, 12]. Yin and Yang use principal component predictive regressions to systematically uncover the components of technical indicators with oil price forecasting power [13]. Liu and Wang *et al.* used moving average rules and macroeconomic indicators to generate density forecasts [14]. They find that the technical indicators generate a more accurate density forecast when compared to the macro variables.

Statistical models are well-known tools for forecasting asset prices. Numerous classes of statistical models, such as random walk [15], generalized autoregressive conditional heteroskedasticity (GARCH) [16], and ARIMA [17] have been used to forecast oil prices. ARIMA is a linear model used for univariate time series analysis and forecasting. Yusof and Rashid *et al.* applied this model to forecast crude oil production in Malaysia for three leading months [18]. Concurrently, Mohammadi and Su used a hybrid ARIMA-GARCH model to model crude oil volatility and compare the accuracy of their framework with four other volatility models [19].

An approach to forecasting has been widely explored in the recent times is the use of machine learning algorithms. Kusonkhum *et al.* used k-Nearest Neighbours (KNN) to predict over-budget construction projects achieving an overall accuracy of 0.86 [20]. Several other machine learning algorithms such as decision trees [21], support vector machines [22], Logistic Regression and Random Forest [23] have been used for prediction and forecasting purposes. Artificial neural networks (ANN) are nonlinear functions that simultaneously capture hidden patterns between input and output variables without any underlying assumptions. Several studies have shown that models based on neural networks have outperformed conventional forecasting and prediction models. ANN is the most commonly used nonlinear AI model. Bakshi *et al.* used convolutional neural network (CNN) model, an extension of ANN, for predicting pregnant shoppers based on their transaction history and purchasing trends [24]. Other extensions, such as the recurrent neural networks (RNN), use loops to iterate over the series while maintaining an internal state that stores information about the steps it has seen so far. These models are efficient in modelling time series data but are often prone to the exploding gradient problem. The LSTM model proposed by Hochreiter and Schmidhuber is a class of RNN models that are not vulnerable to the vanishing gradient problem [25]. The LSTM model has thus been widely adopted for time series modelling as it excels at extracting patterns in an input feature space, where the data spans long sequences. Kubra, Sekeroglu *et al.* used LSTM to perform lung cancer incidence prediction for ten European countries [26]. However, neural networks have their own limitations compared to machine learning algorithms. Premsmith and Ketmaneechairat utilized the logistic regression and Neural Network model for heart disease detection and found that the logistic regression model outperforms the neural network [27].

Recently, several studies have experimented with combining the capabilities of neural networks with the information generated through fundamental, technical, and statistical methods to forecast asset prices. The study indicated that RNN performed better than the two ANN models. Dropsy employed neural networks as a nonlinear forecasting tool for forecasting international equity risk premia in the markets of Germany, Japan, the United Kingdom, and the United States from 1970 to 1990 [28]. Chiroma *et al.* showed that evolutionary neural networks developed using genetic algorithms can show significant performance improvements in predicting crude oil prices compared to known statistical models [29]. The study compiles a comprehensive list of applications of hybrid neural networks to forecast crude oil prices. Wu *et al.* forecasted crude oil prices using a hybrid model utilizing ensemble empirical mode decomposition, comprising sparse Bayesian learning and ARIMA forecasts [30]. Suhermi *et al.* used a hybrid methodology to integrate the ARIMA model and an ANN model and observe an improvement in forecast accuracy compared to the non-hybrid models [31].

The approach that differentiates our work from other existing methods is the use of an ensemble learning method where we train multiple LSTM layers on inputs of crude oil prices along with explanatory variables and statistical forecasts. The novelty is in using an additional ANN that learns from the forecasts of the LSTM networks to yield significantly improved predictions of crude oil prices. The fluctuations in oil prices can be attributed to many factors. Thus, there is a need to incorporate relevant macroeconomic and technical variables to predict crude oil prices accurately. The following section details the explanatory variables used in this study and highlights their relevance in explaining oil prices.

III. Data

We collect crude oil prices from the US Energy Information Administration database, the principal agency responsible for managing energy information. The model is developed for WTI spot price, which is the global standard for crude oil prices. The data is collected for the period starting from 28-07-2000 to 13-05-2019. The macroeconomic indicators and financial time series are collected from the FRED Database, US Energy Information Administration, and Nasdaq Database. Table I presents an overview of the financial time series and their sources. A few of the time series measuring supply and demand levels of crude oil are available only on a weekly scale. In such cases, a linear interpolation is used for filling up the missing value while considering the weekly values on the last Friday of the month. Tables I and II describe the explanatory variables used in this study.

TABLE I. List of Explanatory Variables

Macroeconomic Variable	Frequency	Source
ExxonMobil Closing Price	Daily	Nasdaq Database
Royal Dutch Shell Closing Price	Daily	
Chevron Closing Price	Daily	
PetroChina Closing Price	Daily	
Total Energies Closing prices	Daily	
S&P GSCI Energy Index	Daily	
U.S. Days of Supply of Crude Oil	Weekly	Energy Information Administration database
U.S. Ending Stocks of Crude Oil	Weekly	
U.S. Exports of Crude Oil	Weekly	
U.S. Field Production of Crude Oil	Weekly	
U.S. Imports of Crude Oil	Weekly	
US Refiner Net Input of Crude Oil	Weekly	
Cushing, OK WTI Spot Price	Daily	
Europe Brent Spot Price	Daily	
Cushing Crude Oil Future Contract	Daily	
Henry Hub Natural Gas Spot Price	Daily	
3 Month Treasury	Daily	FRED Database
US Dollar Index	Daily	
Effective Federal Funds Rate	Daily	
Gold Price	Daily	

TABLE II. List of Technical Variables

Technical Indicator Name	Lag length (in days)
Simple moving average (SMA) crossover	1, 3, 6, 9, 12
Moving average convergence divergence (MACD)	(12, 26)
Price Rate of Change (ROC)	3, 6, 9, 12
Relative Strength Index (RSI)	14
Bollinger Bands (Upper, Lower, Squeeze)	20

A. *Fundamental Variables*

Gold is a global commodity generally used to hedge against inflation and has shared a direct relationship with crude oil over time. Wang and Chueh showed that both gold and crude oil prices positively influence each other in the short term [32]. The prices of natural gas, heating oil, and gasoline are used in the study as they serve as significant fuels in the energy mix. Villar and Joutz [33] and Batten *et al.* [34] found a robust leading relationship from natural gas to crude oil arising out of demand and supply factors. Additional variables relating to the import, export, refinery net input, field production, ending stocks, and remaining days of supply of crude oil are also included. These demand and supply factors related to crude oil production are considered to account for the short-term changes associated with production, which are easily affected by several geopolitical and natural factors. Consequently, there is an impact on the reserves and demand for crude oil. The economic policy uncertainty (EPU) index measures the uncertainty in policies related to economic decisions. Crude oil is an asset that is sensitive to policy changes, and the uncertainty index can be used to capture the political environment.

B. *Technical Variables*

While numerous studies conclude the importance of macroeconomic factors in impacting crude oil price movements, a significant strand of literature also considers the importance of technical indicators in explaining price movements. It has been well established that technical indicator can explicitly aid in understanding asset price movements since investors are known to carry out trading decisions based on the technical analysis of historical prices [35, 36]. There are, however, only a few studies that consider both technical and macroeconomic indicators as significant explanatory variables in the context of crude oil [12–14]. This study extends the literature by including relevant trading signals into the forecasting model in addition to the fundamental variables already defined above. The technical indicators chosen are the SMA, RSI, ROC, MACD, and Bollinger Band Squeeze.

The SMA is a momentum-based indicator that generates a buy (sell) signal when a shorter-term (m periods) SMA crosses above (below) the longer-term (n periods) SMA. Ten crossover signals are obtained for crude which are generated using m, n = 1, 3, 6, 9, 12. The RSI is a momentum oscillator that measures the speed and change of asset price movements and is used to identify overbought and oversold conditions. The MACD is a trend following indicator calculated by subtracting the exponential moving average (EMA) over nine periods by the 26 periods EMA. Price ROC is another momentum-based technical indicator that compares the current asset price with the price from a certain number of previous periods. We compute the ROC for 3, 6, 9, and 12 days. Unlike MA and ROC, Bollinger Bands are chart indicators of technical analysis. We use the upper and lower band and the band squeeze in our analysis. These indicators, which depend only on the closing price of the financial asset, are chosen in line with previous studies such as [11, 14, 36, 37]. Crude oil is one of the most actively traded commodities on the exchange, and these indicators could provide some predictive power in determining the price trends for the model.

TABLE III. DESCRIPTIVE STATISTICS OF THE CRUDE OIL PRICES

	Crude Oil Price (in $)
Min	17.5
Max	145.31
Mean	63.1076
Std	26.4272
Skew	0.3322
Kurtosis	-0.7658
Count	4575

Table III describes the summary statistics of the WTI crude oil spot prices. There is a total of 4575 data points. The series has a mean of 63.107 $. Furthermore, a skew of 0.332264 and a kurtosis of −0.765801 exist within the series. Fig. 1 depicts a plot of the crude oil prices, with a vertical line drawn to mark the split between train and test data points, as used by the LSTM networks.

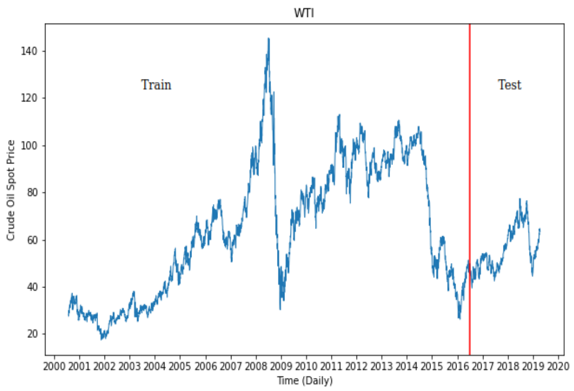

Figure 1. Actual WTI crude oil price from 28-07-2000 to 13-05-2019.

IV. METHODOLOGY

A. Principal Component Analysis (PCA)

We transform the fundamental and technical variables X_i, (i = 1, 2, ...) using PCA [38] to reduce the dimensionality of the data and remove multicollinearity. A defined number of principal components is obtained Z_n, (n = 1, ..., $N \leq I$), which are independent of one another. The PCA technique works by extracting diffusion indexes as a linear combination of the predictors. Working directly with the raw data without standardizing would lead to improper transformations since more weight would be assigned to those variables with relatively higher variances, mainly when the variables are measured in different units. Thus, all variables are standardized before applying the PCA transform for dimensionality reduction.

$$\begin{cases} Z_1 = w_{1,1}X_1 + w_{1,2}X_2 + \ldots + w_{1,t}X_t \\ Z_2 = w_{2,1}X_1 + w_{2,2}X_2 + \ldots + w_{2,t}X_t \\ \quad\quad \vdots \\ \vdots \\ Z_N = w_{N,1}X_1 + w_{N,2}X_2 + \ldots + w_{N,t}X_t \end{cases} \tag{1}$$

where w_i, denotes the jth Eigen value of the i_{th} principal component. The first k, (k<N) principal components that represent a majority of the total information in terms of the variance ratio [8] are selected as inputs to the models.

B. ARIMA Model

The ARIMA model introduced by Box and Jenkins is a statistical technique used for analysing time series forecasting [39]. The ARIMA model uses a linear relationship between the predicted value as a function of a certain number of lagged observations and the lagged values of the residual errors. In general, the ARIMA model is expressed as

$$(1-B)^d Y_t = \mu + \frac{\theta_q(B)}{\emptyset_\varphi(B)} Z_t \tag{2}$$

where,

$$\emptyset_\varphi(B) = 1 - \emptyset_1(B) - \emptyset_2(B^2) - \cdots .. \emptyset_\varphi(B^m) \tag{3}$$

$$\theta_q(B) = 1 - \theta_1(B) - \theta_2(B^2) - \cdots - \theta_q(B^n) \tag{4}$$

$$BY_t = Y_{t-1} \tag{5}$$

Yt denotes the observations, B denotes the backshift operator, and zt denotes the white noise sequence, $a_t W \sim N$ (0, σ^2). ϕ_i (i = 1, 2, ...), θ_j (j = 1, 2, ...,), and μ are model parameters. d denotes the order of differencing.

C. LSTM Model

Traditional feed-forward networks have been extended to develop RNN, which can forecast long data sequences by utilizing internal loops derived from input sequences. RNNs maintain an internal cell state that updates every step of the series, thus remembering and producing results based on past observations. However, RNNs face certain drawbacks due to their inability to factor in errors from older observations while training, making them inefficient to model long-run dependencies. The vanishing (or exploding gradients) problem where weights allocated while training are too small (or large) is one of the major concerns regarding the usability of RNNs. A class of RNN that avoids the issues mentioned above is the LSTM model [25]. It operates with the help of memory cells or states that remember information in the long run and forget past data that is unnecessary.

$$g_t = \sigma(U_g x_t + W_g h_{t-1} + b_f) \tag{6}$$

$$i_t = \sigma(U_i x_t + W_i h_{t-1} + b_i) \tag{7}$$

$$o_t = \sigma(U_o x_t + W_o h_{t-1} + b_o) \tag{8}$$

$$\tilde{c}_t = tanh(U_c x_t + W_c h_{t-1} + b_c) \tag{9}$$

$$c_t = g_t \times c_{t-1} + i_t \times \tilde{c}_t \tag{10}$$

$$h_t = o_t \times \tanh\,(c_t) \tag{11}$$

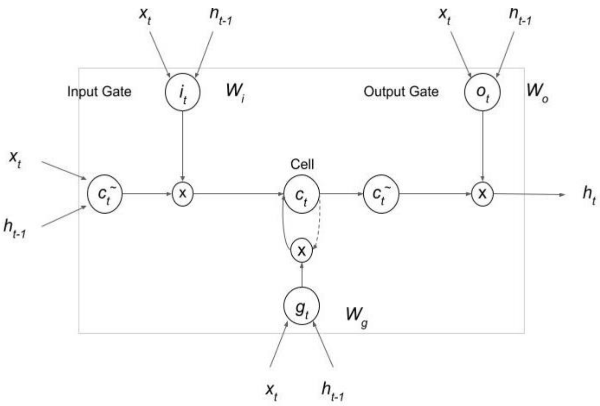

Figure 2. Diagram of LSTM model.

Fig. 2 illustrates the components of the LSTM comprising of a memory cell (c_t) along with three gates: an input gate (i_t), a forget gate (g_t), and an output gate (o_t). The computations associated with the cell state (c_t), the hidden state (h_t), and the three gates are described in Eqs. (6)–(11). $c_0 = 0$ and $h_0 = 0$ are initialized prior to calculations. The input is represented by h_t and the hidden state by $\tilde{c_t}$, at a given time t. The value c_t (input modulate gate) determines the amount of new information received by the cell for every time step. U_0 and W_0 are weight matrices in these equations, b are a bias term, σ is a sigmoid function, $\tan h$ is the hyperbolic tangent function, and the symbol × denotes element-wise multiplication.

D. Test for Comparing Equivalence of Forecast Accuracy

Diebold-Mariano (DM) test and Wilcoxon signed-rank (WS) test are utilized to compare the models' predictive forecasting accuracy. It provides a framework to determine whether the difference in the predictive accuracy of the models is significant for forecasting purposes or is just a result of the choice of data.

- Diebold-Mariano (DM) test

The DM test evaluates each forecast's quality by a predefined loss function g of the forecast error [40]. The null hypothesis of equal predictive accuracy is defined as (z_t) =0, where $Z_t \equiv h(u_{1,t}) - h(u_{2,t})$. The DM statistics are obtained as shown in Eq. (12),

$$DM = \frac{\bar{Z}}{\sqrt{2\pi \hat{g}_z(0)/T}}$$

where, $\bar{Z} = \frac{1}{T}\sum_{t=1}^{T}\left(h(u_{1,t}) - h(u_{2,t})\right)$ and $\hat{g}_z(0)$ is aconsistent estimate of$g_z(0)$.

- Wilcoxon-Signed (WS) Rank Test

The WS Rank test examines whether the difference of forecasting accuracy based on zero-median loss differential is statistically significant. The null hypothesis is given by median (z_t) = 0. If the out-of-sample loss distribution is symmetric, both the tests should give consistent results. WS statistic is given in Eq. (13),

$$WS = \sum_{t=1}^{T} W_+(z_t) rank(|z_t|) \tag{13}$$

where $W_+(z_t) = \begin{cases} 1 & z_t > 0 \\ 0 & otherwise \end{cases}$.

E. Measures of Prediction Errors

The RMSE, MAE), and MAPE are the three-error metrics that have been utilized to evaluate the out-sample forecast of the various prediction's models. Several past works have used these metrics to measure the out-sample efficiency of the training models [41, 42]. In terms of the real and forecasted prices of crude oil, the error metrics are shown in Eqs. (14)–(16),

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_{real,t} - P_{forecast,t})^2} \tag{14}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|P_{real,t} - P_{forecast,t}\right| \tag{15}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|1 - {P_{real,t}}/{P_{forecast,t}}\right| \tag{16}$$

where $P_{real,\,t}$ is the crude oil price, $P_{forecast,\,t}$ is the predicted value for the particular time t, and N is the number of observations.

V. Experimental Setup

The fundamental and technical time-series panel is transformed using PCA, and the first three principal components represents that 80% of the original data are used in this study. The study establishes an LSTM model with a single layer of 64 nodes trained on a single crude oil price input to provide a forecasting capability benchmark. This study investigates how combining the information from fundamental and technical indicators along with ARIMA forecasts into hybrid LSTM models can improve forecasting accuracy. Therefore, the simple LSTM model trained on the historical crude oil prices is used as the baseline for evaluating all proposed hybrid models for the remainder of the analysis. The performance of the models is compared by calculating the RMSE, MAE, and MAPE error measures.

This study proposes two hybrid frameworks to improve the LSTM predictive accuracy. The first model is a multi-input LSTM network (LP) model to evaluate whether information from the chosen explanatory variables can improve prediction accuracy. It involves training the LSTM on the time series of crude oil prices and the PCA transformed series. Three principal components are chosen to cumulatively represent approximately 80% of the variance in the original dataset. The second model builds on the existing model by using information from the predictions of the ARIMA models as an additional input to investigate the advantage of adding statistical forecasts. The ARIMA forecasts are generated using a moving rolling window of 252 days. This model is fit every 252 days to obtain one-step-ahead estimates for the oil prices. Fig. 3 contains a flowchart illustrating the construction of the hybrid model. The ensemble learning-based LPA model has two LSTM models which are trained independently as shown in the figure. The upper LSTM model is identical to the LP model, trained on the PCA and crude oil price inputs while lower LSTM model is trained purely on ARIMA forecasts. Further, ANN with two layers with 256 and 128 nodes, respectively, is used to learn how to best combine the input predictions to make the best output prediction.

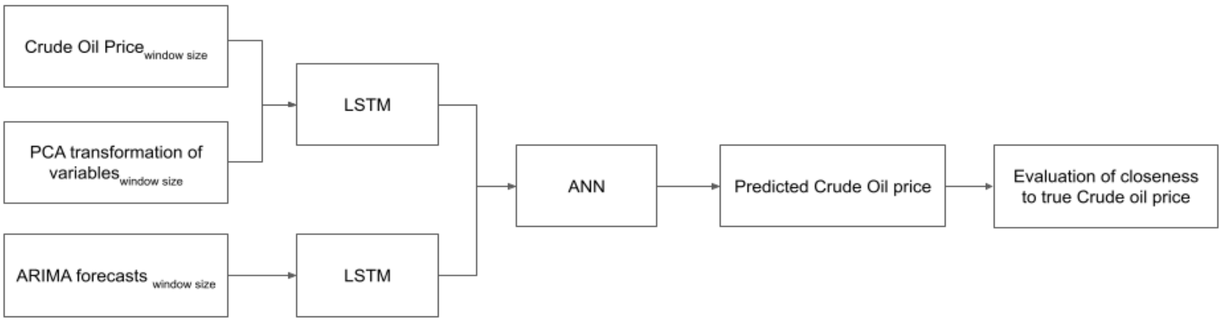

Figure 3. Flowchart of proposed final model.

The proposed models are run with each LSTM and hybrid model repeatedly, trained using inputs taken in four separate rolling windows of sizes 3, 5, 7, and 11 days. Our choice of window sizes is motivated by the periods they represent, where three days represent an intra-weekly period, and 5 and 7 days describe weekly periods. The 11 days' window represents a biweekly time horizon. Each hybrid model was trained twenty-five times independently, and the average of the forecast over all the iterations was used to compute the accuracy. We aim to produce a more reliable and consistent prediction accuracy by averaging all the iterations.

VI. Results and Discussions

This study evaluates the model predictions' errors for the one-day ahead crude oil price forecasts over the out-of-sample period. The following section details the comparisons between the performances of the models. The LSTM model is the baseline for all other proposed hybrid models.

Table IV shows the results for the LSTM, LP, and LPA models. The LP and LPA hybrids outperform the plain LSTM in all the loss functions across all four window sizes. Table V shows the percentage improvement in loss function values for LP and LPA over plain LSTM. An improvement of around 30% is observed when information from the explanatory variables, captured by the PCA transformation, is used as an input into the LSTM to improve the forecast accuracy of crude oil. Notably, the model using a window size of 11-days outperforms other variations, with improvements of 35.19%, 37.72%, and 42.10%, for the RMSE, MAE, and MAPE metrics. Hence, from the results, it is evident that including explanatory variables that account for crude oil's fundamental and technical nature substantially improves the performance of the predictive neural network. It is observed that the improvements for the final hybrid, LPA model are more significant than the LP counterparts across all window sizes and error metrics. For LPA, a window size of 11 days is optimal, showcasing performance improvements of 48.85%, 53.20%, and 50.65% for the RMSE, MAE, and MAPE metrics. The results support the hypothesis that the ARIMA predictions contain additional explanatory information about crude oil price movement, allowing the ensemble learning-based LPA hybrid models to produce better forecasts than the plain LSTM and LP models.

TABLE IV. Error Metrics for Each Proposed Model

Model	WIN_SZ	RMSE	MAE	MAPE
L.5TM	3	1.1943	0.94149	0.01673
	5	1.19075	0.94127	0.01676
	7	1.28716	1.02294	0.01807
	11	1.31064	1.04087	0.01836
LP	3	0.90373	0.74632	0.01223
	5	0.90982	0.7515	0.01226
	7	0.78477	0.61272	0.0101
	11	0.77403	0.58636	0.00969
LPA	3	0.70213	0.51716	0.00945
	5	0.74154	0.55388	0.01002
	7	0.61423	0.44175	0.00822
	11	0.61088	0.44062	0.00826

TABLE V. PERCENTAGE IMPROVEMENT IN LOSS FUNCTION VALUES FOR LP AND LPA OVER PLAIN LSTM

Model	WIN_SZ	RMSE	MAE	MAPE
LP	3	24.333	20.733	26.923
	5	23.823	20.183	26.743
	7	34.293	34.923	39.613
	11	35.193	37.723	42. 103
LPA	3	41.213	45.073	43.543
	5	37.913	41.173	40.093
	7	48.573	53.083	50.843
	11	48.853	53.203	50.653

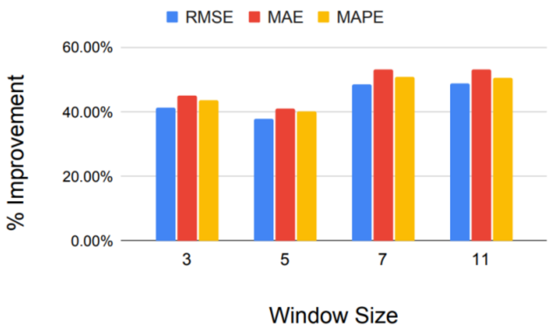

Figure 4. Comparison of prediction errors of LPA to the baseline plain LSTM.

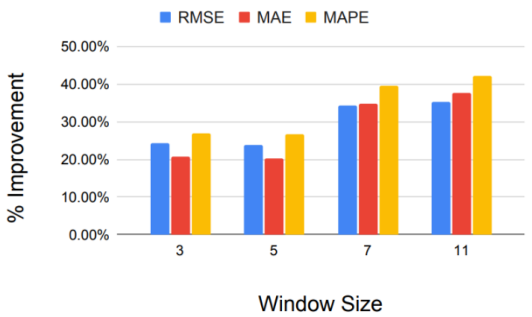

Figure 5. Comparison of prediction errors of LP for baseline plain LSTM.

Figs. 4 and 5 show the improvement in one-step-ahead forecast performance considering the 3-day window size as the benchmark model. Numerous studies have concluded that crude oil prices are unpredictable using traditional econometric methods. The best estimate of future oil prices is the current price itself [43, 44]. However, the relationships between various macroeconomic, geopolitical, supply, demand, and technical factors play a crucial role in explaining oil price behavior [9, 45]. In line with the findings of Miao *et al.*'s [7] and Baumeister and Kilian's [10] research, this study shows a significant confirmation of our hypothesis that the information from traditional forecasting methods further improves price predictions.

Table VI shows the results of the DM and WS tests for equal forecast accuracy. The null hypothesis of these tests suggests that the forecasting models have comparable accuracy, and hence comparisons made between them are not significant. The *p* values obtained from the tests, which involve comparing all model forecasts, are reported in the table. Results associated with the DM tests are reported above the diagonal of each table, while those of the WS tests are reported below the diagonal. At the 95% significance level, all the *p* values suggest that the null hypothesis is rejected, indicating that the out-of-sample forecast accuracy obtained from each model is significantly different from the other.

TABLE VI. DIEBOLD-MARIANO & WILCOXON SIGNED-RANK TEST RESULTS

Model	WIN_SZ	L.5TM	LP	LPA
3	L.5TM		0.00	0.00
	LP	0.00		0.00
	LPA	0.00	0.00	
5	L.5TM		0.00	0.00
	LP	0.00		0.00
	LPA	0.00	0.00	
7	L.5TM		0.00	0.00
	LP	0.00		0.00
	LPA	0.00	0.00	
11	L.5TM		0.00	0.00
	LP	0.00		0.00
	LPA	0.00	0.00	

VII. CONCLUSION

Based on experimental results, both the hybrid models are found to outperform the simple LSTM model. The experimental result showed 20% improvement in the forecasting accuracy for all loss functions for the LP model. The proposed model shown 35.19%, 37.72%, and 42.10% improvement for the RMSE, MAE, and MAPE, respectively by using a window size of 11. Thus, it is evident that adding explanatory fundamental and technical variables helps to improve the forecasting ability of the neural network. The LPA model performs well, and provides 40% improvement for all the error metrics. Models trained on 11 days' window size are again found to be the better performing model. Thus, ARIMA model forecasts add explanatory power to the framework, improving the forecasting accuracy. Hence, this study verified that adding information from fundamental, technical, and statistical methods associated with distinct economic characteristics of crude oil prices as inputs to the proposed LSTM base model significantly improves the one-step-ahead forecasting accuracy of crude oil prices.

Multiple days ahead forecasts of crude oil price movements can be a further extension to this study. This would be pertinent to those with longer investment time horizons. Fluctuations in crude oil prices affect oil producers, governments, oil-dependent industries, and traders. Crude oil also plays a crucial role in the hedging strategies of manufacturers and investors. Crude prices being a commodity of international significance are one of the most important fuel sources and are subject to periods of extreme volatility. The findings of this study may be of relevance to risk management professionals who want to understand the behaviour associated with the crude oil market to avoid potential losses. The results would also aid policymakers concerned with maintaining commodity market stability.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

In this work, Kshitij Kakade, Kshitish Ghate, and Ritika Jaiswal have conceptualized and implemented the idea, while Raj K. Jaiswal has done review, correction and checked for result correctness and validation.

REFERENCES

[1] R. Rapier. (2021). Fossil fuels still supply 84 percent of world energy — And other eye openers from BP's. [Online]. Available: https://www.forbes.com/sites/rrapier/2020/06/20/bp-review-new-highs-in-global-energy-consumption-and-carbon-emissions-in-2019/?sh=1af3546266a1

[2] R. Sari, S. Hammoudeh, and U. Soytas, "Dynamics of oil price, precious metal prices, and exchange rate," *Energy Economics*, vol. 32, no. 2, pp. 351–362, 2010.

[3] Z. Cen and J. Wang, "Crude oil price prediction model with long short-term memory deep learning based on prior knowledge data transfer," *Energy*, vol. 169, pp. 160–171, 2019.

[4] S. T. Katircioglu, K. Sertoglu, M. Candemir, and M. Mercan, "Oil price movements and macroeconomic performance: Evidence from twenty-six OECD countries," *Renewable and Sustainable Energy Reviews*, vol. 44, pp. 257–270, 2015.

[5] J. C. Reboredo and A. Ugolini, "Quantile dependence of oil price movements and stock returns," *Energy Economics*, vol. 54, pp. 33–49, 2016.

[6] Y. J. Zhang and Y. M. Wei, "The dynamic influence of advanced stock market risk on international crude oil returns: An empirical analysis," *Quantitative Finance*, vol. 11, no. 7, pp. 967–978, 2011.

[7] H. Miao, S. Ramchander, T. Wang, and D. Yang, "Influential factors in crude oil price forecasting," *Energy Economics*, vol. 68, pp. 77–88, 2017.

[8] T. Yao and Y. Zhang, "Forecasting crude oil prices with the Google index," *Energy Procedia*, vol. 105, pp. 3772–3776, 2017.

[9] S. Dées, A. Gasteuil, R. Kaufmann, and M. Mann, "Assessing the factors behind oil price changes," European Central Bank, Working Paper Series, no. 855, January 2008.

[10] C. Baumeister and L. Kilian, "Forecasting the real price of oil in a changing world: A forecast combinationapproach," *Journal of Business & Economic Statistics*, vol. 33, no. 3, pp. 338–351, 2015.

[11] L. C. Hsu, "Investment decision making using a combined factor analysis and entropy-based topsis model," *Journal of Business Economics and Management*, vol. 14, no. 3, pp. 448–466, 2013.

[12] S. P. Das, N. S. Achary, and S. Padhy, "Novel hybrid SVMTLBO forecasting model incorporating dimensionality reduction techniques," *Applied Intelligence*, vol. 45, no. 4, pp. 1148–1165, 2016.

[13] L. Yin and Q. Yang, "Predicting the oil prices: Do technical indicators help?" *Energy Economics*, vol. 56, pp. 338–350, 2016.

[14] L. Liu, Y. Wang, and L. Yang, "Predictability of crude oil prices: An investor perspective," *Energy Economics*, vol. 75, pp. 193–205, 2018.

[15] A. Murat and E. Tokat, "Forecasting oil price movements with crack spread futures," *Energy Economics*, vol. 31 no. 1, pp. 85–90, 2009.

[16] Y. Wei, Y. Wang, and D. Huang, "Forecasting crude oil market volatility: Further evidence using GARCH-class models," *Energy Economics*, vol. 32, no. 6, pp. 1477–1484, 2010.

[17] Y. Xiang and X. H. Zhuang, "Application of ARIMA model in short-term prediction of international crude oil price," in *Advanced Materials Research*, Trans Tech Publications Ltd., 2013, vol. 798, pp. 979–982.

[18] N. M. Yusof, R. S. A. Rashid, and Z. Mohamed, "Malaysia crude oil production estimation: An application of ARIMA model," in *Proc. 2010 International Conference on Science and Social Research*, 2010, pp. 1255–1259.

[19] H. Mohammadi and L. Su, "International evidence on crude oil price dynamics: Applications of ARIMA-GARCHmodels," *Energy Economics*, vol. 32, no. 5, pp. 1001–1008, 2010.

[20] W. Kusonkhum, K. Srinavin, N. Leungbootnak, P. Aksorn, and T. Chaitongrat, "Government construction project budget prediction using machine learning," *Journal of Advances in Information Technology*, vol. 13, no. 1, pp. 29–35, Feb. 2022.

[21] A. Alshammari, R. Almalki, and R. Alshammari, "Developing a predictive model of predicting appointment no-show by using machine learning algorithms," *Journal of Advances in Information Technology*, vol. 12, no. 3, pp. 234–239, 2021.

[22] W. Jinxiao, S. Jiaxin, H. Dexin, and Z. Xiaoyu, "Internet financial news and prediction for stock market: An empirical analysis of tourism plate based on LDA and SVM," *Journal of Advances in Information Technology*, vol. 10, no. 3, pp. 95–99, 2019.

[23] T. Daghistani and R. Alshammari, "Comparison of statistical logistic regression and random forest machine learning techniques in predicting diabetes," *Journal of Advances in Information Technology*, vol. 11, no.2, pp. 78–83, 2020.

[24] N. A. Bakshi, P. R. Kolan, and B. Behera, N. Kaushik, and A. M. Ismail, "Predicting pregnant shoppers based on purchase history using deep convolutional neural networks," *Journal of Advances in Information Technology*, vol. 9, no. 4, pp. 110–116, 2018.

[25] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[26] T. Kubra, B. Sekeroglu, and C. Ozkan, "Lung cancer incidence prediction using machine learning algorithms," *Journal of Advances in Information Technology*, vol. 11, no. 2, pp. 91–96, 2020.

[27] J. Premsmith and H. Ketmaneechairat, "A Predictive model for heart disease detection using datamining techniques," *Journal of Advances in Information Technology*, vol. 12, no. 1, pp. 14–20, 2021.

[28] V. Dropsy, "Do macroeconomic factors help in predicting international equity risk premia?: Testing the out-of-sample accuracy of linear and nonlinear forecasts," *Journal of Applied Business Research*, vol. 12, no. 3, p. 120, 2011.

[29] H. Chiroma, S. Abdulkareem, and T. Herawan, "Evolutionary neural network model for west Texas intermediate crude oil price prediction," *Applied Energy*, vol. 142, pp. 266–273, 2015.

[30] Y. X. Wu, Q. B. Wu, and J. Q. Zhu, "Improved EEMD-based crude oil price forecasting using LSTM networks," *Physica A: Statistical Mechanics and its Applications*, vol. 516, pp. 114–124, 2019.

[31] N. Suhermi, D. Suhartono, D. Prastyo, and B. Ali, "Roll motion prediction using a hybrid deep learning and ARIMA model," *Procedia Computer Science*, vol. 144, pp. 251–258, 2018.

[32] Y. S. Wang and Y. L. Chueh, "Dynamic transmission effects between the interest rate, the US dollar, and gold and crude oil prices," *Economic Modelling*, vol. 30, pp. 792–798, 2013.

[33] A. Villar and F. L. Joutz, "The relationship between crude oil and natural gas prices," Energy Information Administration, Office of Oil and Gas, vol. 1, pp. 1–43, 2006.

[34] J. A. Batten, C. Ciner, and B. M. Lucey, "The dynamic linkages between crude oil and natural gas markets," *Energy Economics*, vol. 62, pp. 155–170, 2017.

[35] R. Novy-Marx, "Is momentum really momentum?" *Journal of Financial Economics*, vol. 103, no. 3, pp. 429–453, 2012.

[36] C. J. Neely, D. E. Rapach, J. Tu, and G. Zhou, "Forecasting the equity risk premium: The role of technical indicators," *Management Science*, vol. 60, no. 7, pp. 1772–1791, 2014.

[37] S. Demir, K. Mincev, K. Kok, and N. G. Paterakis, "Introducing technical indicators to electricity price forecasting: A feature engineering study for linear, ensemble, and deep machine learning models," *Applied Sciences*, vol. 10, no. 1, p. 255, 2019.

[38] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics Intell. Lab. Syst.*, vol. 2, no. 1–3, pp. 37–52, 1987.

[39] G. E. P. Box and G. M. Jenkins, *Time Series Analysis Forecasting and Control*, 2nd ed. San Francisco: Holden-Day, 1976.

[40] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *Journal of Business Economic Statistics*, vol. 20, no. 1, pp. 134–144, 2002.

[41] A. Fuertes, M. Izzeldin, and E. Kalotychou, "On forecasting daily stock volatility: The role of intraday information and market conditions," *International Journal of Forecasting*, vol. 25, no. 2, pp. 259–281, 2009.

[42] H. Y. Kim and C. H. Won, "Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models," *Expert Systems with Applications*, vol. 103, pp. 25–37, 2018.

[43] J. D. Hamilton, "Understanding crude oil prices," *The Energy Journal*, vol. 30, no. 2, 2009.

[44] R. Alquist, L. Kilian, and R. J. Vigfusson, "Forecasting the price of oil," *Handbook of Economic Forecasting*, vol. 2, pp. 427–507, 2013.

[45] R. K. Kaufmann, S. Dees, A. Gasteuil, and M. Mann, "Oil prices: The role of refinery utilization, futures markets and non-linearities," *Energy Economics*, vol. 30 no. 5, pp. 2609–2622, 2008.

Kshitij Kakade completed his undergraduate with a master in economics and a bachelor in computer science from BITS Pilani, K.K. Birla Goa Campus. His interests are in the areas of applications of ML in finance.

Kshitish Ghate is a final-year undergraduate student pursuing a master in economics and a bachelor in computer science from BITS Pilani, K.K. Birla Goa Campus. His interests are in the areas of machine learning, natural language processing, financial markets, and behavioral economics with a focus on applications of AI for social good.

Raj K. Jaiswal is working as an assistant professor at SCSDF, National Forensic Sciences University Goa. He has received a Ph.D. degree from the Department of Information Technology, National Institute of Technology Karnataka, Surathkal, India. His research areas of interest include vehicular ad hoc network, SDN/NFV, and cyber security.

Ritika Jaiswal is working as an assistant professor at the Department of Economics & Finance, BITS Pilani, KK Birla Goa Campus. She has done her Ph.D. in the area of commodity derivatives (finance) form the School of Management, National Institute of Technology Karnataka, Surathkal, India. Her research interests include financial economics, market microstructure, risk management and energy finance.