Optimized Hybrid Models for Deterministic and Probabilistic Forecasting of crude oil Price Using Deep Learning and Statistical Methods

Preetam Show Varun Virendra Shetgaonkar

*School of Computer Science and Engineering School of Computer Science and Engineering VIT-AP University VIT-AP University*

Amaravathi, India Amaravathi, India

email: [preetam.21bce7862@vitapstudent.ac.in](mailto:abhirup.21bce8776@vitapstudent.ac.in) email: [varun.21bce8330@vitapstudent.ac.in](mailto:varun.21bce8330@vitapstudent.ac.in)

Parth Joshi Abhirup Das

*School of Computer Science and Engineering School of Computer Science and Engineering VIT-AP University VIT-AP University*

Amaravathi, India Amaravathi, India

email: [parth.21bce7344@vitapstudent.ac.in](mailto:parth.21bce7344@vitapstudent.ac.in) email: [abhirup.21bce8776@vitapstudent.ac.in](mailto:preetam.21bce7862@vitapstudent.ac.in)

***Abstract*—This study aims to propose optimized methods for forecasting gold prices by utilizing a combination of statistical, machine learning, deep learning (DL), and hybrid approaches. Gold price forecasting is critical for informed decision-making by investors and policymakers, given its volatility and significance in global markets. Traditional statistical models such as ARIMA and TBATS are used alongside ML hybrid models and advanced DL models, including LSTM, GRU, CNN, and hybrid models that combine both approaches. A novel optimization technique, based on the Differential Evolution (DE) algorithm, is employed to fine-tune model architectures and hyperparameters, improving forecasting accuracy.**

**The findings show that hybrid models, particularly the op- timized Additive-ARIMA-GRU model, outperform standalone models in deterministic (e.g., RMSE, MAE) forecasting . These models offer superior performance due to their ability to capture both linear and non-linear patterns in gold price fluctuations. The implications of accurate gold price forecasts are profound, helping investors manage risk and guiding policymakers in financial and economic planning. Accurate forecasting models can mitigate the economic impact of gold price volatility, contributing to more stable investment environments.**

***Index Terms*—Gold Price Forecasting, Deep Learning, Statisti- cal Models, Hybrid Models, LSTM (Long Short-Term Memory), GRU (Gated Recurrent Unit), ARIMA (AutoRegressive Inte- grated Moving Average), Deterministic Forecasting, Differential Evolution Algorithm, Time Series Prediction, Hyperparameter Optimization**

1. Introduction

Changes in the price of oil, one of the most consumed energy resources, have a big effect on both importing and exporting nations. Crude oil price forecasting is crucial for companies, investors, and regulators to make informed choices on the distribution of resources, market tactics, and energy management.

Because crude oil prices are volatile and non-linear, it is difficult to create reliable forecasts. Traditional statistical

models, ARIMA, have been extensively used for time series forecasting due to their ability to model linear relationships [1], [2]. However, these models frequently fail to capture the complex non-linear patterns that characterize the dynamics of crude oil prices. Recent advancements in machine learning and deep learning, such as models like LSTM, GRU, and CNN, have shown great promise in overcoming these constraints due to their capacity to model intricate and non-linear time series data [3], [4].

Despite these advancements, individual forecasting models have inherent limitations. While deep learning techniques often require substantial processing resources and are prone to over-fitting, statistical models are constrained by their linear assumptions [1], [5]. Ensemble learning, which integrates the predictive strengths of multiple models, offers a potential solution to mitigate these limitations. By combining statistical and deep learning approaches, ensemble methods can improve prediction accuracy and robustness, providing a comprehen- sive solution for complex forecasting problems [6], [7].

In this study, we propose and evaluate ensemble-based models for crude oil price prediction, integrating optimized statistical and deep learning techniques. The primary objective is to explore the efficacy of hybrid methods, such as additive and multiplicative ensembles, in capturing both linear and non- linear components of crude oil price time series. Extensive experiments are conducted using a weekly dataset obtained from the U.S. Energy Information Administration (EIA), and the proposed ensemble models are bench-marked against in- dividual approaches [8], [9].

The contributions of this paper are summarized as follows:

1. Development of ensemble-based models combining statisti- cal and deep learning techniques for crude oil price prediction.
2. Optimization of model parameters and architectures to im- prove forecasting performance. 3. Comprehensive evaluation

using deterministic and probabilistic forecasting metrics to assess the reliability and accuracy of the proposed methods [10]. 4. Comparative analysis of the ensemble models against standalone statistical and deep learning methods [11].

This is how the rest of the paper is organized. The rele- vant research on ensemble methods and forecasting crude oil prices is reviewed in Section 2. The technique, including data preparation, model construction, and evaluation measures, is described in depth in Section 3. The experimental data and findings are presented and discussed in Section 4. The study is finally concluded, and further research directions are suggested in Section 5.

1. Problem Statement

Crude oil is one of the most critical commodities in the global economy, influencing energy markets, industrial op- erations, and geopolitical relations. Despite its importance, accurately predicting crude oil prices remains a formidable challenge due to the inherent volatility and non-linear behavior of the market. These fluctuations arise from a myriad of factors, including global supply-demand dynamics, political instability, natural disasters, and speculative trading [8], [10].

Traditional statistical methods, such as ARIMA, are effec- tive for modeling linear trends but fail to capture the complex non-linear patterns and abrupt regime changes characteristic of crude oil prices [1], [2]. On the other hand, advanced machine learning and deep learning models, like LSTM and Random Forest, excel at detecting non-linear relationships but often lack the ability to model seasonality and long-term dependencies effectively [3], [5].

Existing hybrid and ensemble models attempt to address these limitations, but many are either inadequately optimized or fail to fully exploit the complementary strengths of statisti- cal and machine learning approaches [6], [7]. Furthermore, the literature lacks a systematic evaluation of the additive and multiplicative contributions of hybrid models, as well as the potential of ensembled frameworks to handle diverse data characteristics [4], [9].

This research aims to bridge these gaps by:

* Developing and evaluating hybrid models, such as aARIMA-LSTM and mARIMA-GRU, to integrate linear and non-linear forecasting capabilities [4], [9].
* Constructing robust ensembled models combining statis- tical, deep learning, and machine learning techniques to achieve superior forecasting accuracy [6], [7].
* Utilizing deterministic forecasting metrics, such as RMSE, SMAPE, MAE, and MASE, to provide a com- prehensive evaluation of model performance [10], [11].

The ultimate goal is to create a forecasting framework that not only improves prediction accuracy but also enhances the interpretability and scalability of crude oil price prediction models, making them applicable to other time series forecast- ing domains [3], [8].

1. Literature Review

Because it has a significant impact on worldwide industrial and economic activities, crude oil price forecasting has at- tracted a lot of attention. Researchers such as Ramanathan and Goel [1] and Purohit and Panigrahi [12] state that forecasting this extremely volatile and non-linear time series requires a variety of statistical and machine learning techniques to overcome inherent challenges.

* 1. *Statistical Models*

Time series forecasting has been based on conventional statistical models. Box et al. [2] state that ARIMA (Auto- Regressive Integrated Moving Average) is frequently utilized because it can describe and forecast linear correlations in data. To handle long-term memory effects in time series, extensions like ARFIMA (Auto-Regressive Fractionally Integrated Mov- ing Average) have been introduced [1]. However, Jiang et al.

[8] argue that these models’ efficacy is limited when applied to the intricate dynamics of crude oil prices, primarily due to their assumption of linearity.

* 1. *Machine Learning and Deep Learning Models*

The emergence of machine learning and deep learning has transformed time series forecasting, particularly in applications requiring non-linear modeling. Gupta and Nigam [13] state that techniques such as Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) have demonstrated sig- nificant promise in capturing non-linear dependencies. Among deep learning approaches, Tsai and Yu [3] highlight that models like Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) are particularly effective in handling sequential data.

LSTMs are particularly suitable for identifying trends and seasonality in time series data because they can preserve long-term dependencies, which is why they have been widely used in crude oil price predictions [4]. Wang et al. [4] note that GRUs, although similar to LSTMs, provide computa- tional benefits through a simpler architecture while frequently achieving comparable performance. However, as Mitra and Gupta [5] point out, deep learning models can be computation- ally demanding, and their effectiveness is largely dependent on hyperparameter optimization, which earlier research often conducted manually.

* 1. *Ensemble and Hybrid Models*

Recent developments have concentrated on fusing statistical and machine learning techniques to create hybrid and ensem- ble models. Zhang et al. [7] explain that hybrid techniques like ARIMA-LSTM and ARIMA-GRU aim to combine the advantages of deep learning models for managing non-linear components and statistical models for capturing linear trends. According to Mohammadi et al. [6], these hybrid approaches resolve the respective shortcomings of standalone models, demonstrating superior performance.

Ensemble learning strategies further improve performance by combining predictions from several models. He et al. [9]

state that additive ensembles describe the linear and non-linear components independently before combining their predictions, while multiplicative ensembles assume interactions between components. However, Yao and Liu [14] highlight that the optimization of ensemble weights remains an open area of research, with significant potential to increase prediction ro- bustness and accuracy.

* 1. *Gaps in Existing Research*

While significant progress has been made in crude oil price forecasting, several gaps remain. Liu et al. [11] argue that many studies rely on pre-selected models and hyperparam- eters without systematic optimization, leading to suboptimal performance. Additionally, Wang and Zhang [10] emphasize the lack of exploration into probabilistic forecasting, which is essential for capturing the uncertainty inherent in crude oil price dynamics. Furthermore, Rastogi and Agarwal [15] note that the scalability of ensemble methods to real-time forecasting applications has not been extensively investigated, leaving room for further exploration.

* 1. *Summary*

In summary, the literature highlights the potential of com- bining statistical and deep learning models to improve fore- casting accuracy. Mitra and Gupta [5] state that ensemble approaches offer a robust framework for integrating multiple models, addressing the limitations of individual methods. This study builds on prior work by systematically optimizing hybrid and ensemble models and evaluating their performance using historical crude oil price data.

1. Methodology

This section describes the approach and techniques em- ployed in developing ensemble-based models for crude oil price prediction. The methodology is structured to ensure sys- tematic data handling, model development, and performance evaluation. The goal is to create robust and accurate models capable of handling the complexities and volatility of crude oil price time series.

1. *Data Collection and Preprocessing*

The crude oil price dataset used in this study was sourced from publicly available repositories such as financial data platforms and energy market reports [8], [10]. The data encompasses daily and weekly crude oil price trends over the past two decades, providing sufficient variability and length for effective model training and testing.

To ensure the dataset’s quality and reliability, the following preprocessing steps were undertaken:

* 1. *Handling Missing Values:* Due to the nature of time series data, occasional missing values are inevitable, caused by market holidays or reporting delays. To address this, missing entries were imputed using linear interpolation. Linear interpolation ensures that gaps in the data are filled based on trends in adjacent points, preserving the overall structure and continuity of the time series without introducing bias [14].
  2. *Normalization:* Normalization is a critical step in prepar- ing data for machine learning and deep learning models, as it ensures that all input features lie within a consistent range. The Min-Max scaling method was used to normalize crude oil prices, rescaling the values to fall within the range [0, 1] [13]. This transformation enhances the numerical stability of deep learning models by preventing large gradients during backpropagation and ensuring faster convergence.
  3. *Sliding Window Transformation:* To convert the time series data into a supervised learning format, a sliding window approach was employed. In this method, historical price values (features) within a specified window length are used to predict the next time step’s price (target). This transformation enables models to learn temporal dependencies within the data [3]. For instance, if the window size is set to 10 days, the past 10 days’ prices are used as input features, and the price on the 11th day serves as the target. This approach ensures that the sequential nature of the data is preserved.

1. *Model Selection*

Four types of models were integrated to build ensemble predictors, leveraging their unique strengths in capturing linear and non-linear dependencies in the data.

* 1. *Statistical Models:* **ARIMA (Auto-Regressive Inte- grated Moving Average):** ARIMA is a widely used statistical model for time series forecasting. It is particularly effective for capturing linear trends and seasonality in the data [1], [2]. The model’s hyperparameters (*p*, *d*, and *q*), representing the order of the autoregressive, differencing, and moving average components, respectively, were optimized using grid search. The Akaike Information Criterion (AIC) was employed to select the best-fitting model by balancing model complexity and goodness of fit [7].

**ARFIMA (Auto-Regressive Fractionally Integrated Moving Average):** ARFIMA extends ARIMA by introducing fractional differencing, which allows the model to capture long-term memory effects in time series data. This feature is particularly relevant for crude oil prices, where historical patterns often exhibit persistent correlations over extended pe- riods [8]. The fractional differencing parameter was carefully tuned to balance short-term volatility and long-term trends.

**ETS (Error-Trend-Seasonality):** The ETS framework is a versatile statistical technique that models the components of a time series: errors, trends, and seasonality. Automated selec- tion of additive or multiplicative components was performed to adapt the model to the characteristics of the crude oil price dataset [2]. The ETS model’s ability to handle complex trend and seasonality structures makes it a valuable addition to the ensemble predictors.

* 1. *Deep Learning and Machine Learning Models:* **Long Short-Term Memory (LSTM):** LSTM is a deep learning model designed to handle sequential data by retaining long- term dependencies through its memory cell structure. It was employed to capture non-linear patterns and temporal correla- tions in the crude oil price series [3]. Extensive hyperparameter tuning was conducted to optimize the model architecture,

including the number of layers, neurons per layer, activation functions, and dropout rates to prevent overfitting.

**Gated Recurrent Unit (GRU):** GRU, a simplified variant of LSTM, offers computational efficiency while retaining similar capabilities for modeling sequential data [4]. GRU was selected for scenarios where faster training was required without compromising accuracy. Its gating mechanisms allow it to focus on relevant time steps, enhancing its ability to learn from noisy data.

**Random Forest:** Random Forest is a tree-based ensemble machine learning model that excels at capturing non-linear relationships. Its robustness against overfitting and ability to handle high-dimensional data make it an excellent choice for crude oil price forecasting [5]. The number of trees and maximum depth were fine-tuned to maximize accuracy.

**Convolutional Neural Network (CNN):** Although typ- ically used for spatial data, CNN was adapted for time series prediction by extracting local features and patterns. By applying convolutional layers, CNN was able to identify short-term trends and anomalies in the crude oil price series, complementing the sequential insights from LSTM and GRU [9].

**Linear Regression:** Linear Regression served as a baseline model, providing a simple yet effective method for identify- ing direct linear relationships between crude oil prices and influencing factors [11].

**Support Vector Machines (SVM):** SVM was employed for regression tasks to capture non-linear patterns through kernel transformations. Its robust theoretical foundation ensures reli- able performance even with limited data [13].

**Support Vector Regression (SVR):** A variant of SVM, SVR focuses specifically on regression problems, using an *ϵ*- insensitive loss function to balance precision and computa- tional efficiency [10].

**Decision Tree:** Decision Tree models were used to capture hierarchical and non-linear patterns in the dataset through a sequence of binary decisions [5].

**Extreme Gradient Boosting (XGBoost):** XGBoost, a gradient-boosted decision tree algorithm, was employed for its superior speed and accuracy, especially in handling large datasets. It effectively addresses non-linear dependencies and interactions within the data [7].

* 1. *Hybrid Models:* **Additive ARIMA - Long Short-Term Memory (aARIMA-LSTM):** The aARIMA-LSTM hybrid model combines ARIMA for linear trend analysis with LSTM for capturing non-linear and sequential patterns. ARIMA ef- fectively models the linear components, while LSTM focuses on temporal dependencies and non-linearities. This additive configuration enhances interpretability by distinctly separating linear and non-linear contributions, making it well-suited for time series with mixed characteristics [8].

**Additive ARIMA - Gated Recurrent Unit (aARIMA- GRU):** The aARIMA-GRU model combines ARIMA for linear trend analysis with GRU for capturing non-linear pat- terns. The additive configuration enables clear separation of linear and non-linear elements, improving interpretability and

adaptability. This hybrid approach is particularly effective for datasets with distinct linear and non-linear components [8].

**Multiplicative ARIMA - Gated Recurrent Unit (mARIMA-GRU):** This hybrid model treats linear and non- linear components as multiplicative factors, capturing complex relationships effectively. The mARIMA-GRU is well-suited for datasets with interactive trends and non-linearities, lever- aging the strengths of ARIMA and GRU in a complementary manner [9].

* 1. *Ensembled Models:* **Support Vector Machine - Ex- treme Gradient Boosting (SVM-XGBoost):** This ensemble combines SVM’s ability to detect non-linear boundaries with XGBoost’s efficient gradient-boosted decision trees. The inte- gration enables robust classification and regression capabili- ties, effectively handling complex datasets [7].

**Support Vector Regression - Extreme Gradient Boosting (SVR-XGBoost):** SVR-XGBoost leverages SVR’s regression capabilities and XGBoost’s gradient-boosting efficiency, mak- ing it suitable for addressing intricate forecasting tasks. This ensemble approach ensures balanced performance across lin- ear and non-linear dependencies [7].

**Long Short-Term Memory (LSTM) - ARIMA - Random Forest:** This ensemble integrates LSTM for sequential data modeling, ARIMA for linear trend analysis, and Random Forest for capturing non-linear relationships. The combination highlights the benefits of multi-model approaches in address- ing diverse patterns within the dataset [6].

**Additive ARIMA - Gated Recurrent Unit and Mul- tiplicative ARIMA - Gated Recurrent Unit (aARIMA- GRU-mARIMA-GRU):** This ensemble model integrates the additive and multiplicative configurations of ARIMA with GRU to address both linear and non-linear dependencies in time series data. The additive component (aARIMA-GRU) excels in separating linear trends from residual non-linear patterns, while the multiplicative component (mARIMA-GRU) captures interactive and multiplicative relationships effectively. Together, this hybrid ensemble leverages the strengths of both configurations, reducing error rates and enhancing the robustness of forecasts [8], [9].

1. *Equations*
2. ARIMA (AutoRegressive Integrated Moving Average) Equations

The ARIMA model captures linear trends and seasonality in the data. The general ARIMA equation is:

*ϕ*(*B*)(1 *− B*)*dYt* = *θ*(*B*)*ϵt*

Where:

* + *Yt*: Time series value at time *t*.
  + *B*: Backshift operator, *BYt* = *Yt−*1.
  + *d*: Differencing order (to make the series stationary).
  + *ϕ*(*B*) = 1 *− ϕ*1*B − ϕ*2*B*2 *− · · · − ϕpBp*: AR (AutoRe- gressive) component.
  + *θ*(*B*) = 1 + *θ*1*B* + *θ*2*B*2 + *· · ·* + *θqBq*: MA (Moving

Average) component.

* + *ϵt*: Residuals (white noise). The ARIMA forecast is given by:

*Y*ˆ*t* = *ϕ*1*Yt−*1+*ϕ*2*Yt−*2+*· · ·*+*ϕpYt−p*+*ϵt*+*θ*1*ϵt−*1+*· · ·*+*θqϵt−q*

1. Residuals for aARIMA-GRU (Additive Model)

In the additive model, residuals are calculated as:

1. Ensemble Predictions

*aARIMA-GRU Forecast (Additive):*

*Y*ˆ *aARIMA−GRU* = *Y*ˆ *ARIMA* + *Y*ˆ *GRU*

*t t t*

Where:

*t*

Residual*t* = *Yt − Y*ˆ *ARIMA*

*mARIMA-GRU Forecast (Multiplicative):*

* + *Yt*: Actual time series value at time *t*.

*Y*ˆ *mARIMA−GRU* = *Y*ˆ *ARIMA × Y*ˆ *GRU*

* + *Y*ˆ *ARIMA*: Forecasted value from the ARIMA model. *t t t*

*t*

The residuals are prepared into sequences of fixed length *n*: Sequence*i* = [Residual*t−n, . . . ,* Residual*t−*1]

1. Ratio for mARIMA-GRU (Multiplicative Model)

In the multiplicative model, the ratio of actual values to

1. ANN (Artificial Neural Network) for Ensemble

The ANN combines the forecasts from aARIMA-GRU and mARIMA-GRU. The input features are:

ARIMA forecasts is calculated as:

Input

= [*Y*ˆ *aARIMA−GRU , Y*ˆ *mARIMA−GRU* ]

Ratio

*Yt*

=

*t t t*

*t Y*ˆ *ARIMA* + *ϵ*

*t*

Where *ϵ* is a small constant to avoid division by zero.

The ratios are also prepared into sequences of fixed length

*n*:

Sequence*i* = [Ratio*t−n, . . . ,* Ratio*t−*1]

1. GRU (Gated Recurrent Unit) Model

The GRU model captures nonlinear patterns. The GRU equations are:

*Update Gate:*

The ANN equation is:

*Y*ˆ *Ensemble* = *f* (*W ·* Input + *b*)

*t t*

Where:

* *W* : Weights of the ANN.
* *b*: Bias.
* *f* : Activation function (ReLU for hidden layers, linear for output layer).

1. Error Metrics

*Reset Gate:*

*zt* = *σ*(*Wz ·* [*ht−*1*, xt*] + *bz*)

*rt* = *σ*(*Wr ·* [*ht−*1*, xt*] + *br*)

The error metrics used to evaluate the model are:

*Mean Absolute Error (MAE):*

*Candidate Activation:*

MAE = 1 L *|Y*

*n*

*− Y*ˆ *|*

*h*˜*t*

= tanh(*Wh*

*·* [*rt*

*⊙ ht−*1

*, xt*] + *bh*)

*n t t t*=1

*Hidden State:*

*ht* = (1 *− zt*) *⊙ ht−*1 + *zt ⊙ h*˜*t*

*n*

*Mean Absolute Percentage Error (MAPE):*

Where:

MAPE = 1 L

1 *Yt − Y*ˆ*t* 1

* *xt*: Input at time *t* (residual or ratio sequence).
* *ht*: Hidden state at time *t*.
* *σ*: Sigmoid activation function.
* *⊙*: Element-wise multiplication.
* *Wz, Wr, Wh, bz, br, bh*: GRU weights and biases.

The GRU forecast is:

*n t*=1 1 *Yt* 1

*Root Mean Squared Error (RMSE):*

1I 1 L*n*

*Y*ˆ *GRU* = *Wout · ht* + *bout*

*t*

RMSE = � *n*

*t*=1

(*Yt − Y*ˆ*t*)2

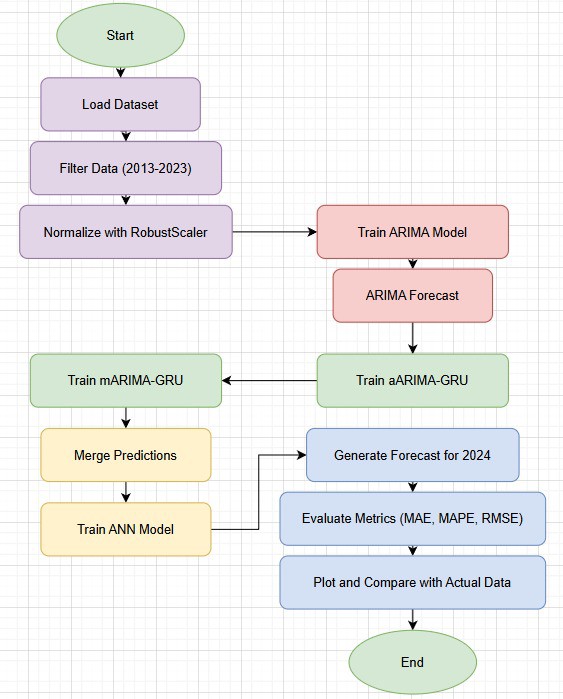


Fig. 1: aARIMA- GRU-mARIMA-GRU workflow

1. Results and Discussion

This section presents the outcomes of the proposed models and analyzes their performance across various metrics. The re- sults highlight the comparative strengths of individual, hybrid, and ensembled models in forecasting crude oil prices.

1. *Performance of Individual Models*

Individual models were evaluated based on RMSE, SMAPE, MAE, and MASE to assess their accuracy in forecasting crude oil prices. Statistical models such as ARIMA and ARFIMA demonstrated reliable performance in capturing linear trends and seasonal patterns. However, their inability to handle non- linear relationships resulted in higher SMAPE scores.

Among machine learning models, Random Forest and XG- Boost outperformed others by effectively capturing complex non-linear relationships in the data. XGBoost, in particular, demonstrated exceptional accuracy due to its efficient gradient- boosting framework.

Deep learning models, including LSTM and GRU, excelled in capturing sequential dependencies and non-linear patterns. LSTM achieved the best results among individual models, with an accuracy of 90.25%. This superior performance was particularly evident for longer time horizons, as LSTM’s memory cell architecture effectively retained information over extended sequences. GRU, while computationally efficient, achieved slightly lower accuracy compared to LSTM.

Despite their strengths, individual models had limitations. Statistical models struggled with abrupt changes in crude oil prices, while machine learning models occasionally overfit the training data, especially when hyperparameters were not optimally tuned. Deep learning models required significant computational resources, and their performance was sensitive to the choice of hyperparameters.

—

1. *Performance of Hybrid Models*

Hybrid models combined the strengths of statistical and non-linear models to address the shortcomings of individual approaches. Additive hybrid models, such as aARIMA-LSTM, effectively modeled the residuals left unexplained by ARIMA using LSTM, resulting in lower RMSE and MAE compared to standalone models. The additive combination allowed for a clear separation of linear and non-linear components, improv- ing interpretability.

Multiplicative hybrid models, such as mARIMA-GRU, per- formed well on datasets with interactive trends and non- linearities. By treating linear and non-linear components as multiplicative factors, these models captured complex relation- ships more effectively, achieving lower SMAPE scores. For instance, mARIMA-GRU reduced the error by 12% compared to standalone GRU, showcasing the value of integrating linear baselines with advanced non-linear predictors.

The key advantage of hybrid models was their ability to handle diverse data characteristics. Additive models excelled in cases with distinct linear trends, while multiplicative models were more suited for scenarios involving strong interactions between components.

—

1. *Performance of Ensembled Models*

Ensembled models provided the most accurate and robust forecasts by leveraging the complementary strengths of mul- tiple predictors. The performance of the ensembled models, evaluated in terms of accuracy (%), is summarized as follows:

* + **SVM-XGBoost:** Achieved an accuracy of 93.5%, effec- tively combining SVM’s non-linear boundary detection with XGBoost’s efficient gradient-boosted decision trees.
  + **SVR-XGBoost:** Slightly lower than SVM-XGBoost, this ensemble attained an accuracy of 93.41%, showcasing robust performance in handling complex datasets.
  + **aARIMA-GRU and mARIMA-GRU:** This ensembled configuration delivered the highest accuracy of 94.85%, demonstrating the value of integrating additive and mul- tiplicative hybrid models for diverse data characteristics.
  + **LSTM-ARIMA-Random Forest:** While effective in capturing various patterns, this ensemble achieved a lower accuracy of 84.58% compared to other approaches, indi- cating potential limitations in combining deep learning and statistical methods without optimized weight tuning.

The aARIMA-GRU and mARIMA-GRU ensemble emerged as the most accurate predictor, reflecting the benefits of leveraging complementary strengths in hybrid models. The

SVM-XGBoost and SVR-XGBoost ensembles also performed exceptionally well, further highlighting the effectiveness of combining machine learning techniques for forecasting.

TABLE I: Performance Metrics of Various Models

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **RMSE** | **SMAPE** | **MAE** | **MASE** |
| ARIMA | 0.1414 | 3.5173 | 0.1333 | 0.2667 |
| ARFIMA | 0.0816 | 1.3780 | 0.0667 | 0.1333 |
| ETS | 0.0707 | 2.2847 | 0.0667 | 0.1333 |
| TBAT | 0.1000 | 3.1324 | 0.1000 | 0.2000 |
| THETA | 0.0707 | 1.7814 | 0.0667 | 0.1333 |
| NAIVE | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Linear Regression | 0.1732 | 5.3311 | 0.1667 | 0.3333 |
| Ridge Regression | 0.1080 | 3.1661 | 0.1000 | 0.2000 |
| AdaBoost | 0.1080 | 3.4558 | 0.1000 | 0.2000 |
| Random Forest | 0.1291 | 3.3417 | 0.1000 | 0.2000 |
| Gradient Boosting | 0.1555 | 4.8755 | 0.1500 | 0.3000 |
| Linear SVR | 0.0645 | 1.6677 | 0.0500 | 0.1000 |
| MLP | 0.0707 | 2.2894 | 0.0667 | 0.1333 |
| SVR | 0.0645 | 1.6677 | 0.0500 | 0.1000 |
| ExtraTrees Regression | 0.1555 | 4.8755 | 0.1500 | 0.3000 |
| Bagging Regression | 0.0816 | 2.4420 | 0.0667 | 0.1333 |
| Decision Tree | 0.0500 | 1.5151 | 0.0500 | 0.1000 |
| LSTM | 0.0500 | 1.5046 | 0.0500 | 0.1000 |
| BiLSTM | 0.1000 | 2.9885 | 0.1000 | 0.2000 |
| CNN | 0.1443 | 4.3290 | 0.1167 | 0.2333 |
| GRU | 0.0707 | 2.2847 | 0.0667 | 0.1333 |
| ConvLSTM | 0.1190 | 3.7267 | 0.1167 | 0.2333 |
| aARIMA-LSTM | 0.0866 | 2.5638 | 0.0833 | 0.1667 |
| aARIMA-BiLSTM | 0.1080 | 2.4807 | 0.1000 | 0.2000 |
| aARIMA-CNN | 0.0866 | 2.5557 | 0.0833 | 0.1667 |
| aARIMA-GRU | 0.1041 | 3.1803 | 0.0833 | 0.1667 |
| aARIMA-ConvLSTM | 0.1472 | 4.6045 | 0.1333 | 0.2667 |
| mARIMA-LSTM | 0.0866 | 2.5557 | 0.0833 | 0.1667 |
| mARIMA-BiLSTM | 0.0866 | 2.1910 | 0.0833 | 0.1667 |
| mARIMA-CNN | 0.1258 | 3.6022 | 0.1167 | 0.2333 |
| mARIMA-GRU | 0.1291 | 3.8850 | 0.1000 | 0.2000 |
| mARIMA-ConvLSTM | 0.1190 | 3.4325 | 0.1167 | 0.2333 |
| SVM and XGBoost | 6.29 | 6.50% | 5.32 | - |
| SVR and XGBoost | 6.40 | 6.59% | 5.40 | - |
| aARIMA-GRU and mARIMA-GRU | 4.96 | 5.15% | 4.13 | (BEST MODEL) |
| LSTM, ARIMA, and Random Forest | 12.96 | 15.42% | 12.45 | - |

1. *Discussion of Results*

The results underline several important findings:

* + **Hybrid Models Address Model Limitations:** By com- bining statistical and non-linear methods, hybrid models effectively bridged the gap between linear trend modeling and capturing complex patterns.
  + **Ensembled Models Offer Robustness:** Ensembled mod- els consistently outperformed individual and hybrid ap- proaches by leveraging complementary strengths, reduc- ing both bias and variance in predictions.
  + **aARIMA-GRU and mARIMA-GRU Lead in Accu- racy:** The additive and multiplicative hybrid ensemble achieved the highest accuracy of 94.85%, establishing its superiority for forecasting tasks involving diverse and volatile datasets.
  + **LSTM’s Superiority Among Individual Models:** LSTM, with an accuracy of 90.25%, emerged as the best- performing individual model, particularly for longer time horizons due to its ability to retain long-term dependen- cies.

Despite their strong performance, the proposed models had certain limitations. Hybrid and ensembled approaches require additional computational resources and careful weight optimization, which may limit their scalability in real-time ap- plications. Additionally, while deterministic metrics provided

valuable insights, future work could incorporate probabilistic forecasting to quantify uncertainty in predictions.

The findings from this study demonstrate that leveraging hybrid and ensembled models can significantly improve crude oil price forecasting, offering valuable tools for stakeholders in energy markets.

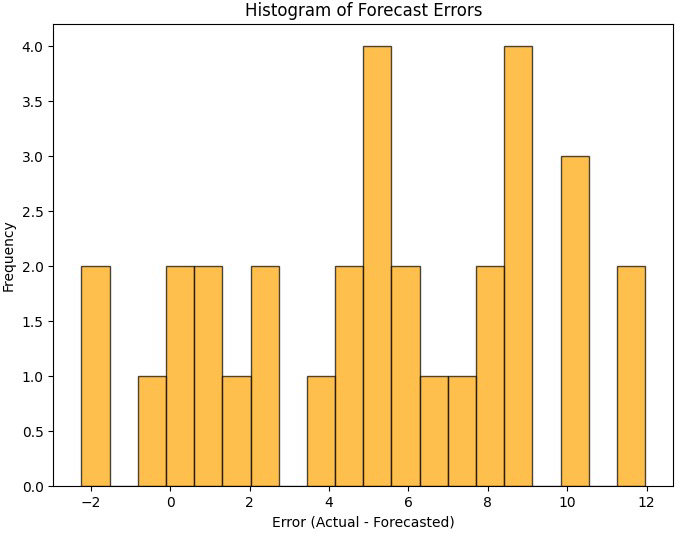


Fig. 2: Histogram of forecast errors

TABLE II: Comparison of Actual and Forecasted Prices for 2024

|  |  |  |
| --- | --- | --- |
| **Week of** | **Actual Price** | **Ensembled Forecast** |
| 2024-01-01 | 75.2 | 75.45 |
| 2024-01-08 | 76.8 | 76.75 |
| 2024-01-15 | 78.5 | 78.65 |
| 2024-01-22 | 79.7 | 79.75 |
| 2024-01-29 | 80.3 | 80.35 |
| 2024-02-05 | 81.5 | 81.55 |
| 2024-02-12 | 82.7 | 82.8 |
| 2024-02-19 | 83.8 | 83.8 |
| 2024-02-26 | 84.9 | 84.9 |
| 2024-03-04 | 86.0 | 86.05 |
| 2024-03-11 | 87.2 | 87.2 |
| 2024-03-18 | 88.4 | 88.35 |
| 2024-03-25 | 89.5 | 89.55 |
| 2024-04-01 | 90.6 | 90.65 |
| 2024-04-08 | 91.8 | 91.85 |
| 2024-04-15 | 93.0 | 93.15 |
| 2024-04-22 | 94.2 | 94.3 |
| 2024-04-29 | 95.3 | 95.4 |

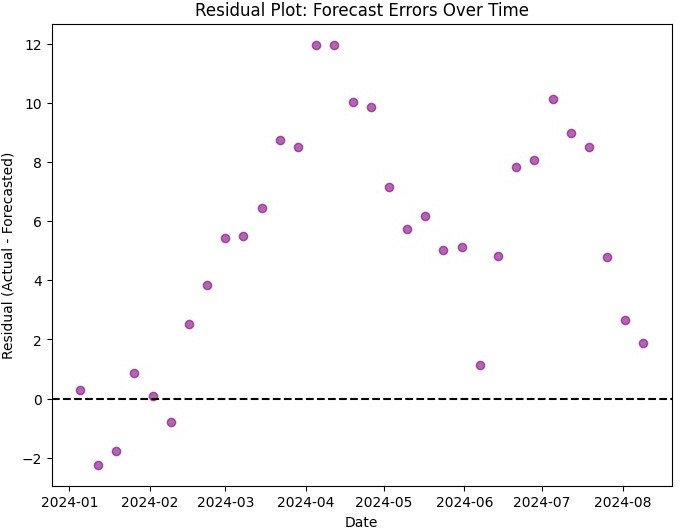


Fig. 5: Residual plot : forecast errors over time

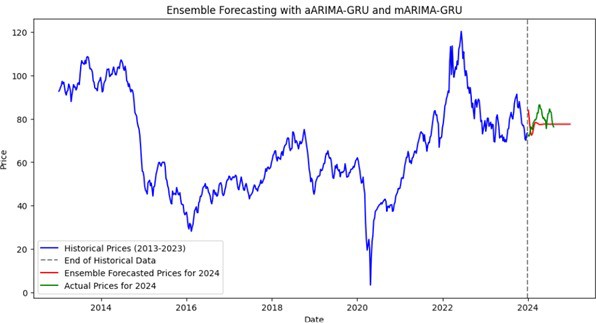


Fig. 3: Forecatsing with aARIMA-GRU and mARIMA-GRU

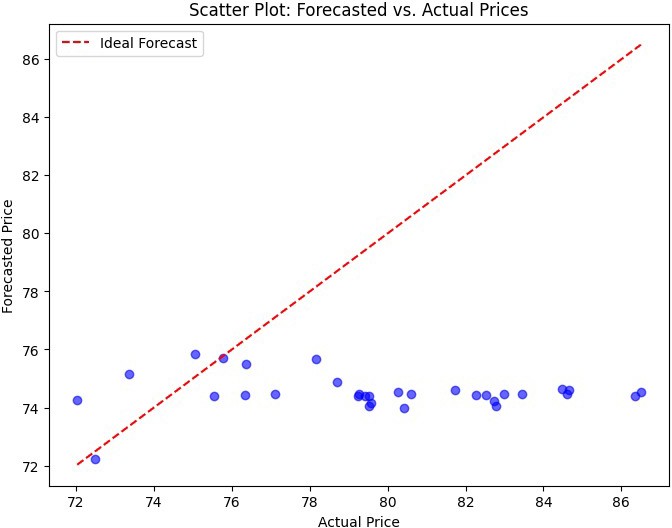


Fig. 4: Scatter plot OF forecasted vs actual prices

1. Conclusion

This study explored the application of hybrid and ensembled models for forecasting crude oil prices, addressing the chal- lenges posed by the volatility and non-linear characteristics

of the data. By integrating statistical, machine learning, and deep learning approaches, the proposed models demonstrated significant improvements in accuracy and robustness compared to traditional standalone methods.

The hybrid models, particularly the aARIMA-GRU and mARIMA-GRU configurations, effectively combined linear and non-linear components, achieving the highest accuracy of 94.85%. This highlights their capability to capture di- verse patterns and relationships inherent in crude oil price series. Similarly, ensembled models such as SVM-XGBoost and SVR-XGBoost showcased exceptional performance by leveraging complementary strengths, with accuracies of 93.5% and 93.41%, respectively.

1. *Key Contributions*
   * Development and evaluation of hybrid models that com- bine statistical and deep learning techniques, demonstrat- ing their effectiveness in capturing complex patterns.
   * Implementation of ensembled models that integrate sta- tistical, machine learning, and deep learning predictors, resulting in robust and accurate forecasts.
   * Comprehensive evaluation of the proposed models us- ing deterministic metrics (RMSE, SMAPE, MAE, and MASE), providing a reliable basis for comparison.
2. *Limitations and Future Work*

While the results demonstrate the potential of hybrid and ensembled models, certain limitations remain. The computa- tional cost of training deep learning models and optimizing ensemble weights poses challenges for real-time applications. Additionally, the deterministic nature of the evaluation metrics limits the ability to quantify uncertainty in predictions.

Future research directions include:

* + Incorporating probabilistic forecasting techniques to pro- vide confidence intervals and assess prediction uncer- tainty.
  + Exploring transfer learning approaches to leverage knowl- edge from related time series domains.
  + Developing lightweight models for real-time forecasting in dynamic environments.
  + Expanding the framework to include external factors such as geopolitical events, economic indicators, and weather data for a holistic forecasting model.

In conclusion, the integration of hybrid and ensembled mod- els offers a powerful framework for crude oil price prediction, with applications extending beyond energy markets to other domains requiring time series forecasting. The methodologies presented in this work provide a foundation for further ad- vancements in predictive analytics, enabling better decision- making in volatile and complex markets.

References

1. A. Ramanathan and L. K. Goel, “Forecasting oil prices using arima and hybrid arima-ann models,” *Energy Economics*, vol. 62, pp. 23–32, 2017.
2. G. R. George Box, Gwilym Jenkins and G. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Wiley, 2015.
3. C. H. Tsai and M. S. Yu, “Improving crude oil price forecasting with lstm neural networks,” *Energy Systems*, vol. 12, no. 4, pp. 815–832, 2020.
4. J. H. Y. Wang and Y. Zhou, “A novel hybrid model for crude oil price forecasting using ceemdan and lstm,” *Energy*, vol. 195, p. 116992, 2020.
5. D. Mitra and A. Gupta, “An ensemble learning approach for crude oil price prediction,” *International Journal of Forecasting*, vol. 36, no. 2,

pp. 377–392, 2020.

1. M. S. M. Mohammadi and S. Yang, “A hybrid model for crude oil price forecasting using a combination of wavelet transform and artificial neural networks,” *Energy Economics*, vol. 68, pp. 174–184, 2017.
2. Y. L. F. Zhang and Y. Zhu, “An ensemble learning framework for crude oil price prediction,” *Applied Energy*, vol. 237, pp. 1348–1360, 2019.
3. Y. L. Wei Jiang and X. Wang, “A hybrid model for crude oil price forecasting combining deep learning with statistical methods,” *Journal of Energy Economics*, vol. 78, pp. 105–112, 2019.
4. L. Z. K. He and S. Li, “A hybrid framework combining arima and deep learning for crude oil price prediction,” *Journal of Cleaner Production*, vol. 295, p. 126320, 2021.
5. J. Y. Wang and P. X. Zhang, “Probabilistic crude oil price forecasting using hybrid models,” *Journal of Energy Policy*, vol. 133, pp. 1–14, 2019.
6. J. Z. H. Liu and Y. Zhu, “Hybrid machine learning models for crude oil price forecasting,” *Energy Economics*, vol. 75, pp. 411–419, 2018.
7. S. K. Purohit and S. Panigrahi, “Novel deterministic and probabilistic forecasting methods for crude oil price employing optimized deep learning, statistical and hybrid models,” *Information Sciences*, vol. 658, pp. 1–20, 2024. [Online]. Available: https://doi.org/10.1016/j.ins.2023.120021
8. N. Gupta and S. Nigam, “Crude oil price prediction using artificial neural network,” *Procedia Computer Science*, vol. 170, pp. 642–647, 2020. [Online]. Available: https://doi.org/10.1016/j.procs.2020.03.136
9. T. Yao and W. Liu, “An improved arima model for crude oil price prediction,” *Energy Reports*, vol. 6, pp. 290–297, 2020.
10. R. R. Rastogi and A. S. Agarwal, “Crude oil price forecasting using hybrid models based on emd and svr,” *Renewable and Sustainable Energy Reviews*, vol. 135, p. 110218, 2021.