University of South Wales
Prifysgol De Cymru

DISSERTATION

Crude oil price forecasting using Deep Learning

My Name: Asim Javed

Registration ID: 30106493

A thesis submitted for the degree of Master of science in

Supervisor Name : Dr Ian Wilson

School of Computer Science and Electronic Engineering

August 2024

Contents

List of Figures

List of Tables

CHAPTER 1

Abstract

Given the volatility of oil prices and their wide ramifications in the economy, the ability to predict the price of crude oil is of paramount importance for businesses, investors, and politicians. This paper compares the standard ARIMA model for crude oil price prediction with the efficacy of two deep learning models, namely CNN and GRU. The paper considered different error metrics like RMSE, MAPE, and R-squared values in the performance evaluation of models. All the models have been trained and tested on a historical pricing dataset. Testing results have shown that the CNN performed better than both GRU and ARIMA, with an R-squared value of 0.84 and MAPE of 14.86 against ARIMA's R squared of -0.57 and MAPE of 11.05. The GRU model also posts competitive results that reveal the superiority of deep learning methods in the detection of sophisticated nonlinear patterns in time series data.

This study underlines the possible benefits of the use of exogenous variables, such as geopolitical events, to increase the accuracy of forecasts. These findings may be of importance to academics and practitioners concerned with managing the complexities of the oil market by making better-informed decisions based on accurate price forecasts.

CHAPTER 2

Introduction

Crude oil price forecasting is a very important problem in both the energy and finance industries due to its great influence on corporate operations and world economies. Accurate forecasts of oil prices can have a great impact on economic policy, supply chain management, and investment strategies. The dynamic and intricate patterns of oil price fluctuations have proven difficult for conventional forecasting techniques like statistical models and econometric methodologies to fully capture. Deep learning, being a section of Artificial Intelligence, is foreseen to further improve accuracy in forecasting [Kumar et al., 2020].
Among others, neural networks are deep learning algorithms that enhance the capability of modeling complex relationships in big datasets. These models learn from underlying patterns and trends from historical oil price data using architectures such as LSTM networks and CNNs, which model both temporal and spatial features. This overcomes some pitfalls of the traditional approaches and hence enables more realistic and detailed predictions [Miao et al., 2017]. Given that a wide range of factors, such as geopolitical developments, market mood, and economic indicators, affect crude oil prices, deep learning offers a viable means of enhancing forecasting precision and assisting well-informed decision-making in the energy industry.

Chapter 3

Background

Energy is undoubtedly the driving force of human civilization. Because of this fact, energy-related problems always exist in the world. For example, even though scientists have already discovered new sources of energy over the past few decades, according to the "BP Statistical Review of World Energy June 2016," fossil fuels such as coal, crude oil, and gas still dominated the majority of energy consumption at 85.9 percent. Apart from crude oil consumption at 32.9 percent outshines that of gas at 23.8, and coal of 29.2 [Amoco, 2001]. Therefore, the panics always happen in the world economy due to the high volatility of crude oil price. Generally, a sharp rise in the price of crude oil would drive oil-importing nations into recession and inflation, which would further hurt the world economy. The fast and steady economic growth thus heavily depends on the precision of the crude oil price forecast [Gholamian et al., 2005]. Following first and second Gulf Wars, the severe high increase of the price in crude oil destroyed severely the world's economy. It is noticed that other factors such as supply and demand in crude oil market, the stock market, US dollar exchange rate, Speculative trade of Derivative Contracts of oil, Natural calamities- the earthquakes occurred in Ecuador in 1987, etc enhance the market risks [He et al., 2010]. While the demand for oil depends on a range of economic factors, it also actually depends on the price of oil. Supply is determined by the competitive behavior of non-OPEC producers and the decisions of OPEC producers in negotiations. The interaction between supply and demand, with ever-changing and unpredictable market conditions, makes the task of structural modeling hard [Dees et al., 2007].

Generally, the price of crude oil is one of the major variables that really affects world financial markets. Since oil prices may influence everything from bond and stock markets, investors and academics studying the oil market have always given keen interest to providing a reliable estimate of the price of oil and its variation [Wan and Kao, 2015]. For time series models, the forecasting process is carried out utilizing past data. Both linear and nonlinear models are used in the process. However, determining whether a correlation is linear or nonlinear appears to be a difficult process, and it is sometimes asserted that it is difficult to determine if a correlation is fully linear or nonlinear [Safari and Davallou, 2018].

Crude oil is one of the most critical commodities within the world's economy, affecting various streams in manufacturing, transportation, and even energy generation. In order for stakeholders such as firms, investors, and regulators to mitigate risks and make highly informed decisions, the crude oil price forecast needs to be accurate. Historically used traditional methods include econometric models and statistical techniques. However, the complex and nonlinear structures in the historical crude oil price data are often too challenging for these methods to capture [Li et al., 2024].

[Hochreiter and Schmidhuber, 1997]. Because these models can learn from enormous datasets and recognize complex patterns, they are a good fit for predicting jobs that are beyond the scope of traditional approaches [LeCun et al., 2015].

Crude oil affects many different areas of the global economy, making it a vital component. Crude oil is essential for driving industry, transportation, and the production of electricity. As of 2022, it made up around 33 percent of the worlds primary energy use. The stability and expansion of the global economy are strongly impacted by the price of crude oil [Hamilton, 2009]. For example, changes in oil prices have been associated with increases in inflation; a 10 percent increase in oil costs might result in an increase in inflation of 0.4 percent in developed economies and 0.8 percent in underdeveloped nations. Furthermore, the transportation industry is largely reliant on crude oil; in fact, transportation fuels account for around 55 percent of global oil consumption. Because of this dependence, fluctuations in oil prices can have a direct impact on the price of products and services, which can have an impact on consumer spending and corporate operations [Yu et al., 2016]. Changes in the price of oil have an impact on the world commerce market as well. Oil exports comprised around 8.5 percent of all global trade in 2021, demonstrating the significance of oil in international trade. The volatility of the price of oil also affects the financial markets. Economic recessions and stock market alterations have historically followed large shocks to oil prices, as seen by

the 1973 oil crisis and the 2008 financial crisis [Zhang et al., 2015]. These figures highlight how much the price of crude oil affects the world economy, highlighting the need for precise forecasting models to reduce risks and guide policy choices [Plourde and Watkins, 1994].
Another finding of their increasing use is in time series forecasting. CNNs, first created for image processing, may now be used to extract features from the time series data by identifying regional patterns and trends. Hybrid models that blend CNNs with RNNs or LSTMs exploit the advantages of both architectures, leading to an improvement in the performance of the prediction. The hybrid models, while having the capability to handle both the temporal and geographical representation of time series data, form a very strong basis for crude oil price prediction [Chen et al., 2016].
Since 1970, there have been several fluctuations in the price of crude oil on a worldwide scale, which has had a significant effect on society and the economy [Hagen, 1994]. First, supply and demand act mainly on the price of crude oil, just as it does with any other commodity. In addition, many other factors such as weather, stock market performance, economic development, political problems, psychological anticipation, and even sudden events have dramatic impacts. Due to the aforementioned variables, the global crude oil market is highly volatile and nonlinear with dynamic volatility and high irregularity [Yu et al., 2008]. Due to these variables, the global crude oil market is highly volatile and characterized by complex nonlinearity, dynamic volatility, and high irregularity.
While deep learning has its revolution, forecasting crude oil price is still a challenging task due to its intrinsic volatility and sensitivity to market speculations, economic indicators, and geopolitical events. Therefore, continuous innovation and research in model designs and training methods are required to raise prediction accuracy and dependability. Other exogenous variables that could increase the prediction potential of these models further include natural disasters, political stability, and technology breakthroughs.
In order to investigate the use of deep learning techniques in crude oil price forecasting, this study will assess the efficacy of several model architectures, such as RNNs, LSTMs, GRUs, and hybrid models. Our goal is to help players make more informed decisions in the unpredictable oil market by utilizing deep learning to produce forecasts that are more accurate and dependable.

3.1 Research Aim

It develops and evaluates an accurate and reliable deep learning model that can be effectively utilized in crude oil price forecasting. It assembles ways research may improve the level of accuracy regarding crude oil price predictions by using cutting-edge deep learning approaches such as CNN, GRU, and ARIMA. This will be achieved through very accurate and timely projections that will enhance the quality of decisions made by stakeholders, which, in turn, will increase economic stability and reduce the risks associated with the fluctuation in the prices of crude oil. The study also investigates how changes in other external factors-for instance, different worldwide events or economic data-affect the price of crude oil.

3.2 Research Objectives

Some of the research objectives for this topic are:

- Build a Deep Learning Model: Construct a complex deep learning model that is specialized for predicting crude oil prices, employing Gated Recurrent Unit and Convolutional Neural Networks (CNN).
- Evaluate Model Performance: By contrasting the constructed model with conventional statistics and machine learning techniques, determine the model's resilience and accuracy.
- Include External Factors: To improve the models prediction power, include important external factors like economic data and wars.
- Provide Precise and Timely estimates: Assist stakeholders in their decision-making processes by providing accurate and timely estimates of the price of crude oil.

3.3 Research Motivations

The research on deep learning-based forecasting of crude oil prices is inspired by the great impact of the prices on the world economy. As a key energy supply, crude oil influences the world's economic stability, production costs, and inflation. The main risk related to volatility may be decreased with proper price projections that help governments, corporations, and investors make better decisions. Most of the time, the conventional method of forecasting fails to capture such an intricate nonlinear pattern in the swings of oil price. Deep learning has emerged as a viable answer to this problem because of its high degree of sophistication for processing data and identifying patterns. This project will hence apply deep learning methods to enhance forecast accuracy by providing handy tools and insights to support decision-making in a highly dynamic and important market.

3.4 Research Questions

Some of the research questions for this topic are:

- How can the performance of these deep learning models be improved by optimizing hyperparameters?
- In terms of precision and robustness, how do deep learning models (CNN and GRU) for crude oil price predictions do against more conventional models (ARIMA)?
- Which outside variables such as geopolitical events or economic data have the most effects on the price of crude oil, and how can these variables be successfully included into deep learning models?

3.5 Ethical Consideration

Several ethical issues need to be taken into account while doing research on deep learning-based crude oil price predictions in order to maintain the study's integrity and social effect. First and foremost, confidentiality and privacy of data are crucial. Ensuring the privacy of all parties involved requires that all data utilized be publicly accessible and anonymised. To ensure that sensitive data is safeguarded during the study process, strict data security measures must be put in place to prevent unauthorized access.

When reporting findings, honesty and transparency are essential. Accurate presentation of data, transparent methodology, and candid discussion of any limits are all requirements for researchers. The scientific community gains credibility and confidence as a result of this transparency, which makes peer review and validation possible.

Ensuring fairness in the model requires addressing potential biases. To find and reduce any biases that can affect the predictions, testing and continuous evaluation should be done. By doing this, the model is guaranteed to produce fair and accurate projections.

It is also morally required to take stakeholders' wider effects into account. To prevent unwarranted fear or misinterpretation that might have negative impact on markets and communities, researchers must be mindful of the economic and social ramifications of their projections and appropriately convey their results.

CHAPTER 4

Literature Review

Most of the research into prediction methods has been driven by the economic importance and volatility of crude oil prices. Some common econometric models for time series forecasting include the Vector Autoregression (VAR), the Generalized Autoregressive Conditional Heteroskedasticity (GARCH), and ARIMA (Autoregressive Integrated Moving Average) models [Adebiyi et al., 2014]. But these models have faced difficulties due to the complicated and non-linear nature of crude oil price fluctuations, which has prompted academics to look into more sophisticated methods [Sadorsky, 2006].

Deep learning models have become more effective tools for time series forecasting in recent years because of their capacity to identify complex patterns and non-linear correlations in data [Zhang et al., 1998]. Intrinsic, time series data is sequential; LSTMs are a variation of RNNs very capable of processing sequential data. Furthermore, the architecture of LSTMs solves the vanishing gradient problem in RNNs, thus the ability of LSTMs to remember long-term dependencies and increase their prediction accuracy[Hochreiter and Schmidhuber, 1997].

Other architectures that have been used to make forecasts for oil price variations include CNNs and hybrid models which combine CNNs with LSTMs. CNNs adapt to time series forecasting through the identification of local patterns and trends in data. They are generally used in image processing [Borovykh et al., 2017]. The combined advantages of CNNs and LSTMs allow hybrid models to better capture temporal and spatial relationships in the data [Kim and Kim, 2019].

The higher effectiveness of deep learning models in predicting crude oil prices has been shown in a number of studies. To anticipate West Texas Intermediate (WTI) crude oil prices, for example, [Kurdi, 2022] used an LSTM model and discovered that it performed better in terms of prediction accuracy than conventional models. In a similar vein, author study, which contrasted LSTM with other machine learning models such as Support Vector Machines (SVM), discovered that LSTM was superior at identifying the seasonality and underlying patterns in crude oil prices [Chen et al., 2018]. Crude oil price dynamics is highly susceptible to exogenous variables such as macroeconomic statistics, market sentiment, and even geopolitical events. Moreover, it has also been found that deep learning models do better in forecasting when these external factors are taken into consideration. For example, the author improved prediction accuracy and robustness by adding news sentiment analysis and macroeconomic variables to an LSTM model [Abdollahi and Ebrahimi, 2020].

Xiong, Bao, and co-authors have proposed an updated hybrid model based on the EMD feed-forward neural network modeling framework. In the new model, SBM was also integrated to provide an advanced capturing of the sophisticated dynamics in the prices of crude oil. While multi-step-ahead forecasting, according to the author, represents a method of extrapolating the crude oil price series for which no available outputs are available in the interest horizon, one-step-ahead prediction provides no insight into long-term future crude oil price behaviors. They also suggested that multi-step-ahead crude oil price forecasts should be used more frequently by practitioners and government agencies in decision-making about investments that are linked with the oil sector, because it is of greater value to decision-makers compared to one-step-ahead forecasts [Xiong et al., 2013]. Author and others share the belief that multi-step-ahead forecasts, which accurately represent the future dynamic behavior of crude oil prices, may assist practitioners and governmental organizations in making and adjusting various decisions for various time periods [Yu et al., 2015].

Data fluctuation networks predictive model (DFNPM), a unique method to crude oil price forecasting based on data fluctuation network, was suggested by Wang et al. and outperforms the conventional prediction models in terms of both direction and level of prediction [Wang et al., 2018]. Even if the results of the research that are cited demonstrate the benefits of artificial intelligence techniques for nonlinear and extremely complicated financial data, these techniques still have drawbacks such as over-fitting and parameter sensitivity. They can quickly reach a local minimum and cause more issues [Kulkarni and Haidar, 2009]. Author used GARCH-M models to forecast various sub-sequences and EMD to convert the

denoised crude oil price into several sub-sequences. Comparing these merged models to other benchmark models, the prediction power is higher [Shambulingappa et al., 2020].
In order to anticipate the daily price of crude oil, Elshendy et al. gathered four predictors that were taken from Wikipedia, Twitter, Google Trends, and the Global Data on Events, Location, and Tone (GDELT) database. Compared to earlier research, we place greater emphasis on a collaborative examination of media sources and Google Trends [Elshendy et al., 2018]. A recent study's authors examined over 45 million news messages to demonstrate the potent forecasting value of internet news messages for oil prices. They maintained that text mining algorithms can extract useful information from online crude oil news and that severe occurrences (such political unrest and economic growth) have a significant impact on the oil market [Wex et al., 2013].
Tang et al. proposed a new ensemble learning approach that incorporates extended extreme learning machine (EELM) with complementary ensemble empirical mode decomposition (CEEMD) to enhance the accuracy of crude oil price forecasting [Tang et al., 2015]. In contrast to sophisticated combining techniques (where the weights rely on the covariance matrix of the forecasting error), Hendry and Clements revealed that a straightforward combining approach, simple averaging, displayed good functioning [Guidolin and Timmermann, 2009]. According to Terasvirta, the use of both the linear and nonlinear models combined in forecasting generates better results than making use of just one nonlinear model. Combining the two forecast models may enable both forecasting models to improve the capability of providing forecasts about the future. Furthermore, this will combine different forecasting models with each other so that it would capture a range of linkages as well as patterns present in time series data[Stock and Watson, 2006].

4.1 Research Gap

Although deep learning approaches have made significant progress in forecasting crude oil prices, there are still a number of unresolved research gaps. First off, even if LSTM and hybrid models have outperformed more conventional techniques in terms of accuracy, their capabilities may still be expanded by adding additional real-time and varied data sources. Existing models frequently overlook the possible influence of high-frequency trading data, social media sentiment, and real-time geopolitical developments in favor of historical price data and a small number of economic variables [Zhang et al., 1998].

Another big issue in deep learning models is their interpretability. Because these models have a black-box character, they can hardly be applied in practice for financial decision-making, even though they might learn complicated patterns in data, since the underlying decision-making process cannot be comprehended. Techniques that provide more interpretability and transparency for these models have to be developed if they are to see wider adoption [Samek et al., 2021].

Furthermore, there is uncertainty about the robustness and applicability of the deep learning models in various market scenarios and time frames. Various studies done so far have often been focused on specific datasets and time periods, which may not reveal the full range of market dynamics. For confirmation of consistency and its reliability, these models need to be extensively tested under variant conditions [Kurdi, 2022].

CHAPTER 5

Risk and Mitigation

In order to ensure effective decision-making, stakeholders must handle certain risks associated with utilizing deep learning models to estimate crude oil prices. Model overfitting, in which a model performs well on past data but badly on unknown data, is a serious concern. It is crucial to use strategies like regularization, cross validation, and a large dataset with a range of market situations to improve generalization in order to reduce this risk.
Further, unpredictability with exogenous factors, such as geopolitical events resulting in dramatic changes of oil price, brings yet another dimension of concern. There might be discrepancies between forecasts and their actual prices coming from such unpredictable causes. In this respect, a state-of-the-art development toward responsiveness such a forecast might include analytics performed using live data, further enriched by sentiment analyses supportive for identification of market feel and trend identification.
Furthermore, if a single model like ARIMA is used and it is unable to adjust to changes in the market, stakeholders may be put at risk. Accuracy and robustness can be increased by using an ensemble technique that incorporates predictions from several models, such as CNN and GRU.

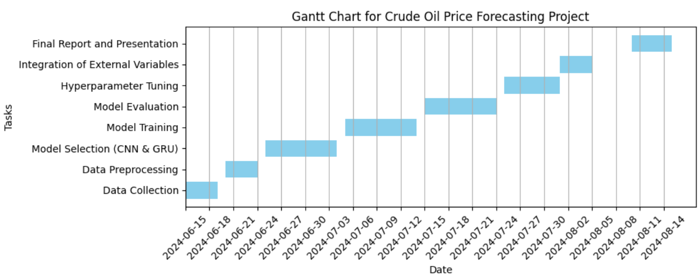

Figure 5.1: Gantt Chart

Above diagram 5.1 shows for the Crude Oil Price Forecasting Project, a Gantt chart is shown. The timeline and order of the several tasks: data collection, preprocessing, model selection, training, assessment, integration of external variables, hyperparameter adjustment, and final report and presentation are shown in the chart. It is anticipated that the project will be finished by August 14, 2024.

CHAPTER 6

Data and its Analysis

This data set from Yahoo Inc. contains historical behaviour of crude oil prices from 23rd August 2000 to 10th August 2024, supplemented by daily trading volumes and adjusted closing. In total, there are 24 years of huge data, which might be necessary for market analysis, understanding volatility, or spotting a trend. Data preprocessing-standardization, development of features by filling in missed values-will definitely help in modeling more accurately. The visual representation of this using line charts and heatmaps indicates patterns and relations within the data; thus, the correlation between prices will also be understood. These stages are of great importance in the energy sector for good prediction models and making wise investment decisions.

The dataset of the crude oil price is taken from [Yahoo Inc., 2023]. The information provided includes the historical prices of crude oil from August 23, 2000, to August 10, 2024. Each entry consists of the daily trade data that includes trading volumes, adjusted closing prices, along with opening, highest, lowest, and closing prices. The essential characteristics are: "Open," "High," "Low," "Close," "Adj Close," "Volume," representing the dynamics of the market and activities of investors day in and day out. This in-depth dataset gives way to studying in-depth the behavior, turbulence, and tendencies of the dynamics within crude oil prices within a 24-year period. It is, therefore, possible to make proper investment in choices and strategy in energy-based industries if the data required to construct a number of predictive models from the use of machine learning techniques is available.

6.1 Preprocessing of Data

Data preprocessing is necessary for preparing crude oil prices for analysis and modeling. Missing values in the dataset have been handled with the use of several methods, including imputation and deletion of data, in the intent of cleaning it up. More importantly, with respect to machine learning algorithms, normalizing or standardizing price values ensures that each feature has equal contribution in analysis [García et al., 2015]. Furthermore, you may improve model performance by converting date-time formats and developing pertinent time-based features like moving averages. These preparation procedures guarantee data quality and raise prediction model accuracy [Witten et al., 2005]. Some of the preprocessing tasks that is performed are:

- Checking whether there is any presence of null values or not in the dataset. Also, reset index is applied to the Date column so that it is treated as a column.
- Splitting of dataset into training,testing and validation. From 2000 to 2018 training dataset is used, 2018 to 2022 validation dataset is used and from 2022 testing dataset is used.

6.2 Data Visualization

Data visualization is the representation of information to find patterns and trends in data visually, which is an important first step in examining data. Hence, the interaction of the trends in the data, in this case the historical pattern of crude oil prices and their associated volatility, often shows up most clearly in line plots, bar plots, and scatterplots. All this is possible with the use of visualization in order to identify outliers, understand how data is distributed, and present complicated information in an understandable way. Visualization enhances decision-making by putting numerical data into a clear, visual context that enables the better interpretation and effective communication of results [Few, 2013].

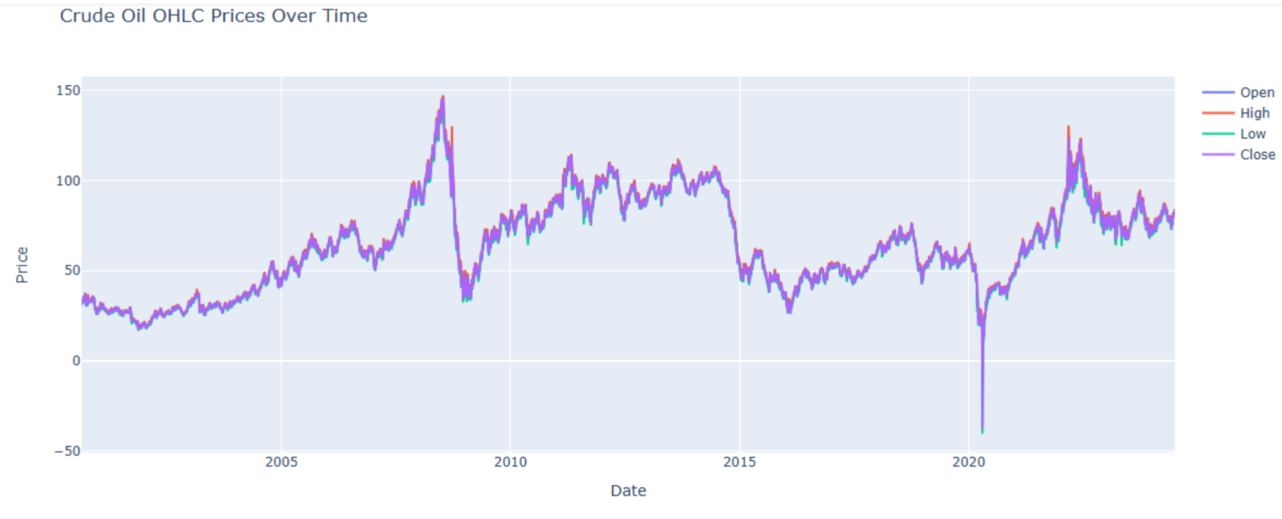

Figure 6.1: Crude Oil OHLC Price over time

Above diagram 6.1 displays the crude oil prices throughout time at the Open, High, Close, and Low points. In general, prices rose between 2004 and 2008 before reverting to their starting points. After that, they rose once again, reaching a high somewhere around 2014, and then decreased until 2020. The prices have increased since then. The graph illustrates the overall volatility of crude oil prices, which are impacted by a number of variables such as world events and economic situations.

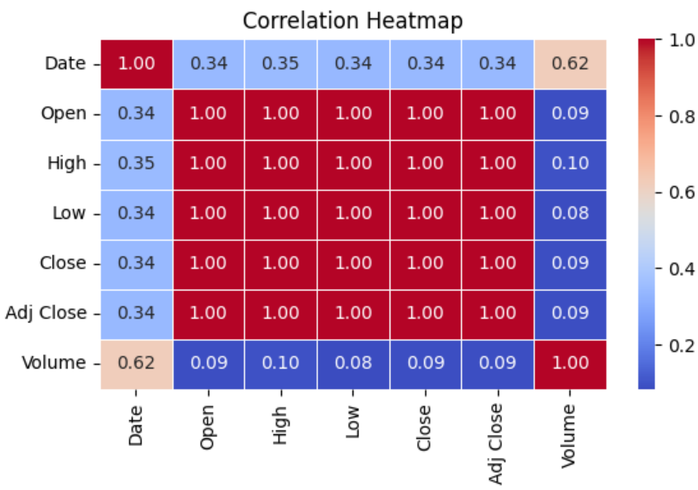

Figure 6.2: Heatmap

Above diagram 6.2 shows the correlation between the various data characteristics is displayed on the heatmap. Red denotes a strong positive connection, blue denotes a strong negative correlation, and white indicates no association. The color intensity signifies the degree of the correlation. As we can see, there is a strong positive connection between the characteristics Open, High, Low, Close, and Adj Close, a mild positive correlation between the features Date and Volume, and a negative correlation between the features.

6.3 Summary of Data

summary	open	high	low	close	Adj close
count	6050	6050	6050	6050	6050
mean	64.50	65.49	63.43	64.49	64.49
Std	25.28	25.55	24.98	25.29	25.29
min	-14.00	13.69	40.32	-37.63	-37.63
25 percent	45.30	46.22	44.34	45.33	45.33
50 percent	62.99	63.80	62.11	63.09	63.09
75 percent	83.26	84.54	82	83.26	83.26
max	145.19	147.27	143.22	145.28	145.28

Table 6.1: Summary of the Data for Crude Oil Price

The table 6.1 shows important information on the distribution and fluctuation of prices throughout the studied period may be found in the statistical summary of the crude oil price dataset. 6,050 observations make up the dataset, which represents the daily trade data.
The average price of crude oil during this period remained constant, staying at around 64.50 to 65.49 dollars, as suggested by the mean values of "Open," "High," "Low," "Close," and "Adj Close" prices. The standard deviation figures coming to almost 25 dollars show the high degree of price volatility because of the change in market circumstances over this period of time.
The minimum and maximum values may show extreme variances: the lowest "Open" price ever recorded was -14.00 dollars, while the greatest "High" price was 147.27 dollars. Such a huge difference gives evidence of extraordinary market volatility, probably brought on by rare economic or geopolitical events in general. Quantiles Q1 (25th percentile) prices are about 45.30 dollars, whereas the 75th percentile prices, Q3, are approximately 83.26 dollars. These quartile figures give more detail on the spread of price. It shows that the center of this dataset tends to, where half of the trading days closed within this range of prices.
Overall, the summary figures show a wide range of volatile crude oil prices during the studied period, underscoring the need for reliable forecasting models to successfully navigate

such a dynamic market.

CHAPTER 7

Methodology

The ARIMA and CNN models are two important techniques in data analysis and time series forecasting. By making use of convolutional layers, activation functions, and pooling operations, a CNN is capable of finding out the spatial hierarchies from grid-like data like photographs. On the other hand, moving averages, autoregression, and differencing are important in the analysis and forecasting of time series data in ARIMA models. This section reviews some working theories, model architectures, and real-world applications of these techniques, along with their benefits and drawbacks in managing intricate datasets and enhancing prediction accuracy across a wide range of applications.

7.1 Gated Recurrent Unit

One kind of recurrent neural network (RNN) architecture called the Gated Recurrent Unit (GRU) was created to solve the vanishing gradient issue in conventional RNNs. Gating mechanisms are used by the GRU to regulate information flow. The update gate and the reset gate are the two main gates of the GRU [Cho et al., 2014]. The update gate establishes the amount of prior data that must be sent to the following stage. The reset gate determines the amount of historical data to be erased. By removing unnecessary data and preserving long term relationships, these gates assist the model in preserving crucial context across sequences

[Ravanelli et al., 2018]. The equations of this topic:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \tag{7.1.1}$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \tag{7.1.2}$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \tag{7.1.3}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \tag{7.1.4}$$

where:

- z_t is the update gate at time step t,
- r_t is the reset gate at time step t,
- $\tilde{h}_t$ is the candidate activation at time step t,
- h_t is the hidden state at time step t,
- σ is the sigmoid function,
- $\tanh$ is the hyperbolic tangent function,
- W_z, W_r, and W are weight matrices,
- x_t is the input at time step t,
- h_{t-1} is the hidden state from the previous time step.

7.1.1 Working Principle

In the current time step, the update gate decides how much of the prior concealed state should be carried over. How much historical data should be erased is determined by the reset gate. The reset gate modifies the prior hidden state, which is used to compute the candidate activation. The current input is also used [Chung et al., 2014]. With the update gate controlling the process, the final hidden state is a linear interpolation between the candidate activation and the prior hidden state. GRUs can preserve long-term dependencies because to these methods, which also make them appropriate for time-series prediction and language modeling applications [Dey and Salem, 2017].

The full gated unit comes in a number of variants, with gating executed in different ways utilizing the bias and the prior concealed state, as well as a more straightforward version known as the basic gated unit.

7.1.2 Implementation of Model

The model consists of three GRU layers and one Dense layer. The first GRU layer returns sequences with 256 units, the second returns sequences with 128 units, while the third does not return sequences with 64 units. The last Dense layer produces a single value.

The Huber loss function, which is resistant to outliers, and the RMSprop optimizer with a learning rate of 0.001 are used to create the model. After that, the model is trained for 10 epochs with a batch size of 32 and a validation split of 20 percent using the training datasets X_train_data and y_train_data.

Predictions for the validation set are produced by the model after training. The scaler is then used to inversely translate these predictions to match the original scale, guaranteeing a comparable difference between the forecasts and the real values. Consistency is checked by printing the shapes of the real values and the forecasts.

The model predictions on the test set are also inverted back to the original scale. This procedure is done to ensure that the performance of the model may be fairly compared with the actual test data. This should enable the model to better identify complex temporal patterns in the data.

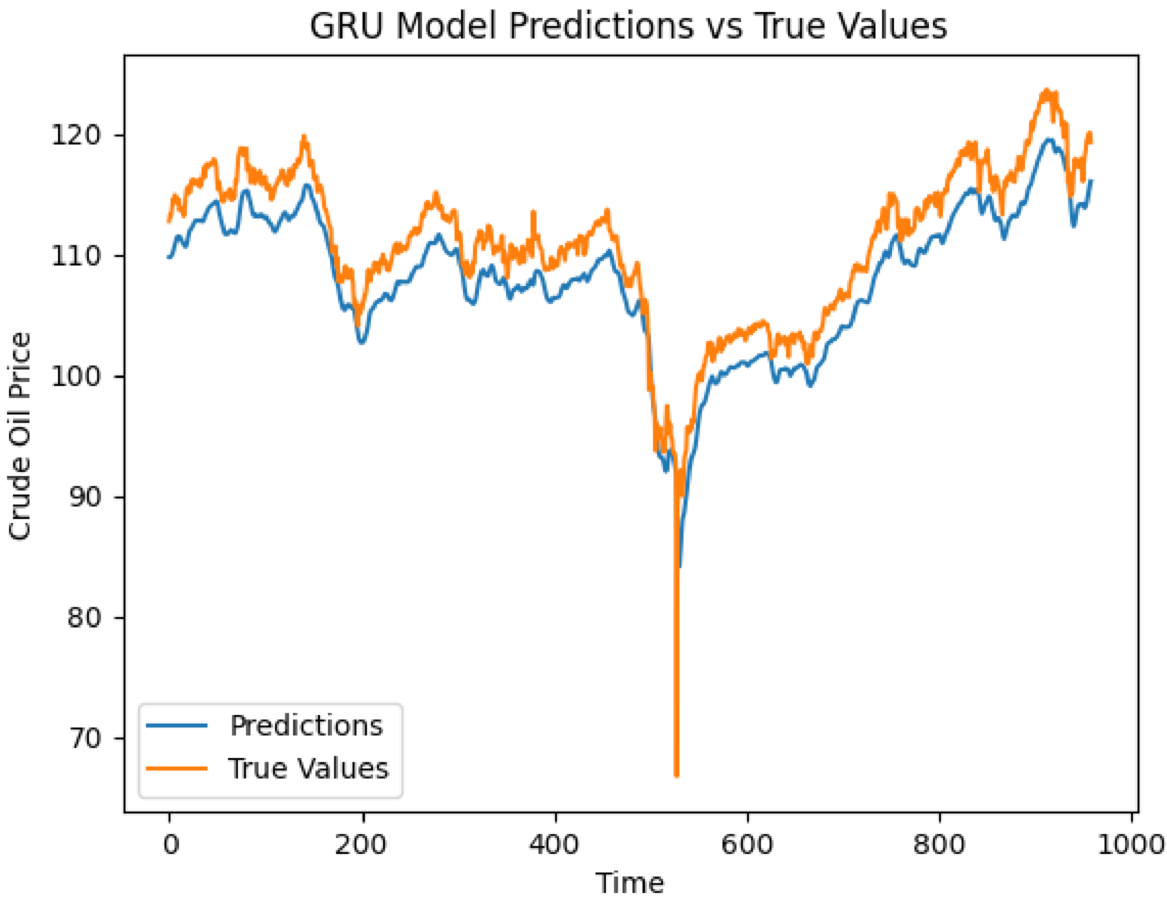

Figure 7.1: Gru Model Predictions vs True Value(Validation)

Above diagram 7.1 plots the real values of the crude oil price against the predictions of a GRU model. The orange line is ground truth, and the blue is the model's prediction. Though the model follows the general trend in the market, it fails to capture the sudden crash and subsequent bounce of the crude oil prices occurring at the 500th time step. This happens because of the global event covid 19 when the crude oil price market value goes to negative. This suggests more model optimization can be done in order to estimate the sudden and rapid changes arising in the oil market.

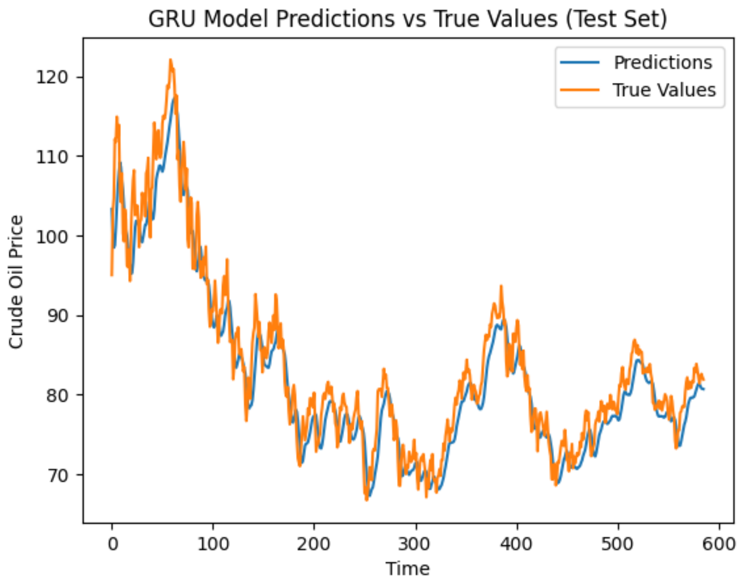

Figure 7.2: Gru Model Predictions vs True Value(Test Set)

Above diagram 7.2 displays a GRU model's crude oil price predictions compared to the test sets actual values. The forecasts and real values are closely aligned, indicating a well performing model. Although there are various peaks and troughs in the model, overall the broad trend of the time series is captured. The GRU model's ability to predict crude oil prices properly is demonstrated by the close match between forecasts and true data.

7.1.3 Advantages and Disadvantages

Some of the advantages of the model are:

- Ability to Capture Complex Patterns: As a kind of recurrent neural network, GRU models are particularly good at capturing complex patterns and temporal dependencies in time-series data. This makes them a good fit for forecasting the price of crude oil, which is subject to a variety of time-dependent variables.
- Handling of Long Sequences: Unlike conventional RNNs, GRUs are able to handle long sequences without experiencing the vanishing gradient problem. This feature is especially helpful for projecting crude oil prices, because precise forecasting sometimes depends on previous data covering lengthy time periods.
- Efficiency: GRUs are computationally more efficient while maintaining the capacity to learn long-term dependencies, as they have fewer parameters than LSTMs because they

do not have a dedicated memory cell. When working with the huge datasets common to financial forecasting, this efficiency might be vital.

Some of the disadvantages of the model are:

- Data Requirements: In order to train deep learning models including GRUs to attain high accuracy, a significant amount of data is needed. Obtaining and preparing large amounts of historical data on crude oil prices may be difficult and time-consuming.
- Overfitting: GRU models are susceptible to overfitting the training set, particularly in cases when the model's complexity is high or the training set is small. For forecasting erratic markets like crude oil, overfitting can lead to poor generalization to unknown data.

7.2 Convolutional Neural Network

A Convolutional Neural Network-commonly referred to as CNN-is a deep learning model intended for the analysis of data with a grid-like architecture, such as photographs. Multiple layers make up a CNN, where each layer acts in a distinct manner in a step-by-step process of representing and learning features hierarchically in the input data [Venkatesan and Li, 2017].

- Convolutional Layers: These layers recognize local patterns like edges, textures, and forms by applying convolutional filters to the input data. Every filter produces feature maps that emphasize the existence of particular patterns by convolving over the input [Balas et al., 2020].
- Activation Function: To enable the network to learn intricate patterns, an activation function (sometimes called ReLU) adds non-linearity after each convolution [Aghdam et al., 2017].
- Pooling Layers: Using techniques like max pooling, pooling layers downsample feature maps to cut down on spatial dimensions without losing crucial information. This stage aids in the reduction of overfitting and computational complexity.
- Fully Connected Layers: At the end, fully connected layers use the acquired features to execute sophisticated logic and generate forecasts. Like in typical neural networks, these layers link every neuron in one layer to every other layer's neuron.

- Training and Backpropagation: CNNs are trained using backpropagation, in which the model iteratively improves its performance by adjusting its weights depending on the gradient of the loss function.

Convolution Operation

The convolution operation can be defined as:

$$(S * W)(i, j) = \sum_m \sum_n S(i - m, j - n)W(m, n) \tag{7.2.1}$$

where S is the input, W is the filter (kernel), and $*$ denotes the convolution operation.

Activation Function

The Rectified Linear Unit (ReLU) activation function is defined as:

$$f(x) = \max(0, x) \tag{7.2.2}$$

Pooling Operation

Max pooling operation is defined as:

$$P(i, j) = \max_{m,n}(S(i + m, j + n)) \tag{7.2.3}$$

where P is the pooled feature map.

Fully Connected Layer

The output of a fully connected layer can be defined as:

$$y = f(Wx + b) \tag{7.2.4}$$

where W is the weight matrix, x is the input vector, b is the bias vector, and f is the activation function.

Loss Function

The loss function, for instance, the cross-entropy loss for classification, is defined as:

$$L = -\sum_i y_i \log(\hat{y}_i) \tag{7.2.5}$$

where y_i is the true label and $\hat{y}_i$ is the predicted probability.

Backpropagation

The gradient of the loss with respect to the weights is calculated as:

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial W} \tag{7.2.6}$$

7.2.1 Working Principle and Model Architecture

The CNN is a deep learning model that intends to take the picture of the data in a grid-like format and analyze them. Architecture: In the CNN, a multitude of layers are present that automatically select and adaptively determine the spatial hierarchies of different characteristics from the input picture. In detail, it is the following architecture and working principle:

7.2.2 Working Principle

During the process of the operation of a CNN, there is always an activation function that comes after the convolution operations in every stage; usually, the ReLU is used. A convolution procedure produces feature maps out of input data by using a series of filters, which are sometimes referred to as kernels. This training procedure yields these filters, which help in the detection of deeper layer elements such as textures, edges, and more intricate patterns [Valueva et al., 2020].

Then, some pooling layers like max pooling are performed in order to shrink the spatial dimensions of feature maps and keep the most important information with reduced computational cost. Since a series of convolution and pooling procedures have been performed through a number of layers, the network can learn abstract and sophisticated properties in a progressive manner[Tsantekidis et al., 2017].

7.2.3 Model Architecture

1. **Input Layer**: The input to the network is an image with dimensions $32 \times 32 \times 3$, representing a color image with height, width, and three color channels (RGB).

2. **Convolutional Layer 1**:

    - Apply 32 filters (kernels) of size 3×3.
    - Convolution operation:

    $$(S * W)(i, j) = \sum_m \sum_n S(i - m, j - n)W(m, n)$$

    where S is the input image, W is the filter, and $*$ denotes the convolution operation.

    - Activation function: ReLU

    $$f(x) = \max(0, x)$$

3. **Pooling Layer 1**:

    - Apply max pooling with a 2×2 filter and a stride of 2.
    - Max pooling operation:

    $$P(i, j) = \max_{m,n}(S(i + m, j + n))$$

4. **Convolutional Layer 2**:

    - Apply 64 filters of size 3×3.
    - Convolution operation:

    $$(S * W)(i, j) = \sum_m \sum_n S(i - m, j - n)W(m, n)$$

    - Activation function: ReLU

    $$f(x) = \max(0, x)$$

5. **Pooling Layer 2**:

    - Apply max pooling with a 2×2 filter and a stride of 2.
    - Max pooling operation:

    $$P(i, j) = \max_{m,n}(S(i + m, j + n))$$

6. **Fully Connected Layer**:

   - Flatten the output from the previous layer to a vector.
   - Apply a dense (fully connected) layer with 128 units.
   - Activation function: ReLU

$$f(x) = \max(0, x)$$

7. **Output Layer**:

   - Apply a dense layer with 10 units (for 10-class classification, e.g., digits 0-9).
   - Activation function: Softmax

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

7.2.4 Implementation of Model

The Keras Sequential API is used to build the CNN model. Included in the model architecture are:

- **Convolutional Layers**: 32 filters, a 3 kernel size, and the ReLU activation function are present in the first Conv1D layer. 64 filter, 3 kernel size, and ReLU activation function are features of the second Conv1D layer.
- **Layer Pooling**: For dimensionality reduction, two MaxPooling1D layers with a pool size of two are used.
- **Flattening Layer**: To transform the 2D matrix data into a 1D vector, use a flatten layer.
- Fully Connected Layers: a 50 unit dense layer with ReLU activation. To forecast the desired value, an output Dense layer of 1 unit is used.

It trains a model using the RMSprop optimizer, a learning rate of 0.001, and the Huber loss function. This model will train on the training set in batches of size 64, with 10 percent for the validation split across ten epochs. It will make predictions after being trained on the test and validation sets. Afterwards, these predictions are inverted back to return the data to its original scale. To ensure that they agree, the forms of the actual values and the predictions are printed last.

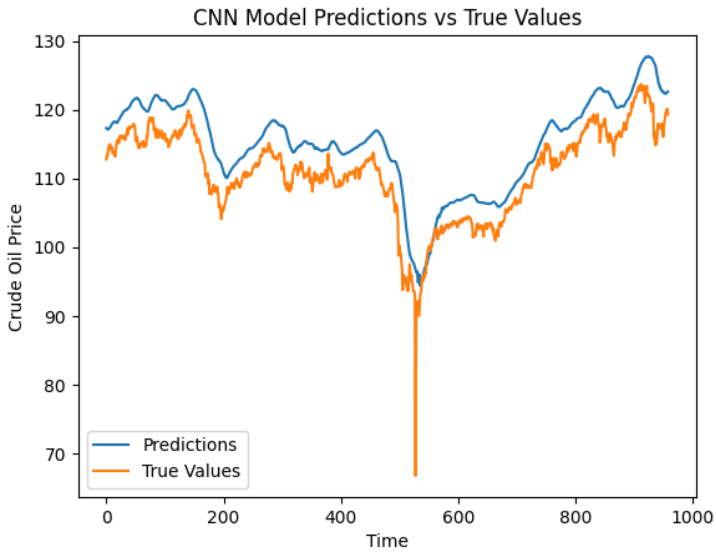

Figure 7.3: CNN Model Predictions vs True Value(Validation Set)

Above diagram 7.3 displays the actual values of crude oil prices over time in comparison to the predictions made by a CNN model. While the model appears to be able to accurately anticipate the overall trend in oil prices, its forecasting becomes less precise when the price is changing quickly. The dramatic decline in oil prices around time 500 is especially difficult for the model to handle. The model performs very well overall, while there is room for improvement in terms of accuracy and ability to handle abrupt price swings.

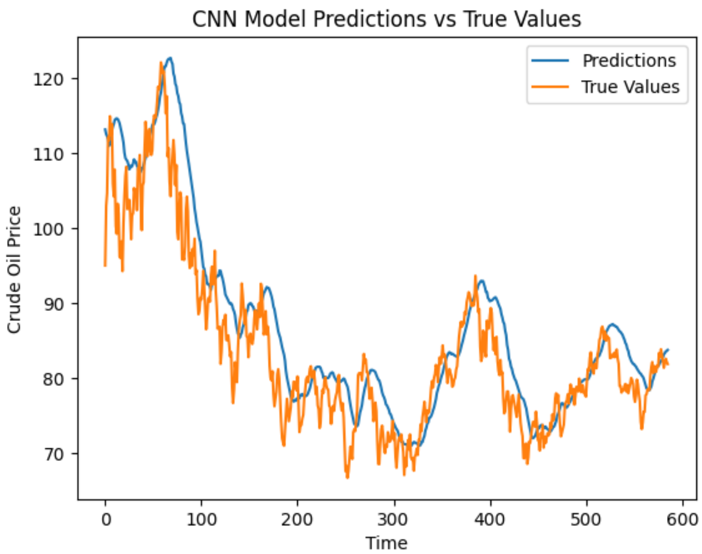

Figure 7.4: CNN Model Predictions vs True Value(Test Set)

Above diagram 7.4 displays the CNN model's predictions for the price of crude oil in relation to the actual values. The orange line shows the actual numbers, while the blue line shows the predictions made by the model. The model has significant variations, especially near the peaks and troughs, but overall it tracks the genuine values' trend. This implies that while the model may have trouble forecasting the precise time and size of price variations, it may be able to identify broad trends in the data. All things considered, it appears that the model comprehends the fundamental trends in the data on crude oil prices.

7.2.5 Advantages and Disadvantages

Some of the advantages of this model are:

- Feature Extraction:From unprocessed input data, CNNs automatically identify and extract pertinent characteristics. Time series data, such as crude oil prices, benefit greatly from this as intricate patterns may be found without the need for manual feature engineering.
- Effectiveness:CNNs are computationally efficient because of parameter sharing and their hierarchical structure, which is especially useful when working with big datasets. In comparison to certain other deep learning models, this enables quicker training and inference.

Some of the disadvantages of this model are:

- Complexity:CNN architectures can have several layers and hyperparameters that require adjustment, making them complicated structures. Smaller businesses or projects may find it difficult to handle this complexity since it demands a great deal of experience and computer power.
- Overfitting: Overfitting is a risk that CNNs may face, particularly when trained on noisy or limited datasets. Overfitting can result in poor generalization and erroneous predictions on unseen data, which is a key issue in financial forecasting.

7.3 Auto Regression Integrated Moving Average

When it comes to time series forecasting, one widely used statistical technique is the ARIMA (AutoRegressive Integrated Moving Average) model. Moving average (MA), differencing (I), and autoregression (AR) are its three constituent parts.

- The dependence between one observation and many lag observations is used in the AR portion.
- To make the time series stationary, the I portion entails differencing the raw observations.
- When a moving average model is applied to lagged observations, the link between an observation and a residual error is modeled in the MA section.

The ARIMA (AutoRegressive Integrated Moving Average) model is represented as:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{7.3.1}$$

where Y_t is the time series value at time t, ϕ are the coefficients of the AR part, θ are the coefficients of the MA part, and ε_t is white noise.

7.3.1 Working Principle

Capturing temporal interdependence in a time series dataset is the foundation upon which the ARIMA (AutoRegressive Integrated Moving Average) model functions. Autoregression (AR), differencing (I), and moving average (MA) are its three main parts.

- The autoregression (AR) component represents the correlation between an observation and a predetermined number of delayed observations, or earlier points in time. According to the theory, values from the past shape values now.
- Integrated (I): To attain stationarity, which denotes that the statistical characteristics of the series hold steady throughout time, differencing is used to the time series data. Because ARIMA is predicated on a stationary process, this is vital.

- Moving Average (MA): When a moving average model is applied to lagged observations, this component simulates the connection between an observation and a residual error.

$$AutoRegressive : Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t \tag{7.3.2}$$

where Y_t is the current value, ϕ represents the coefficients of the AR terms, and ε_t is white noise.

2.

$$Y_t' = Y_t - Y_{t-1} \tag{7.3.3}$$

where Y_t' is the differenced series.

3. **Moving Average (MA)**

$$Y_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{7.3.4}$$

where μ is the mean of the series, θ represents the coefficients of the MA terms, and ε is white noise.

7.3.2 Auto Correlation and Partial Auto Correlation

The association between a time series and its own historical values is measured by autocorrelation. It facilitates the discovery of trends and connections in the data at various delays. The autocorrelation function (ACF) at lag k is defined as:

$$\rho(k) = \frac{\text{Cov}(Y_t, Y_{t-k})}{\sigma^2} \tag{7.3.5}$$

where Y_t is the value of the time series at time t, Y_{t-k} is the value at lag k, Cov denotes covariance, and σ^2 is the variance of the series.

After accounting for the influence of intermediate delays, partial autocorrelation calculates the correlation between a time series and its lag values. It is helpful in determining an autoregressive model's proper order.The partial autocorrelation function (PACF) at lag k is defined as:

$$\phi(k) = \text{corr}(Y_t, Y_{t-k} | Y_{t-1}, Y_{t-2}, \ldots, Y_{t-(k-1)}) \tag{7.3.6}$$

where $\phi(k)$ is the partial autocorrelation at lag k and corr denotes the correlation after controlling for the values of the intervening lags.

Both ACF and PACF are essential tools in time series analysis for identifying the order of ARIMA models and understanding the structure of the data.

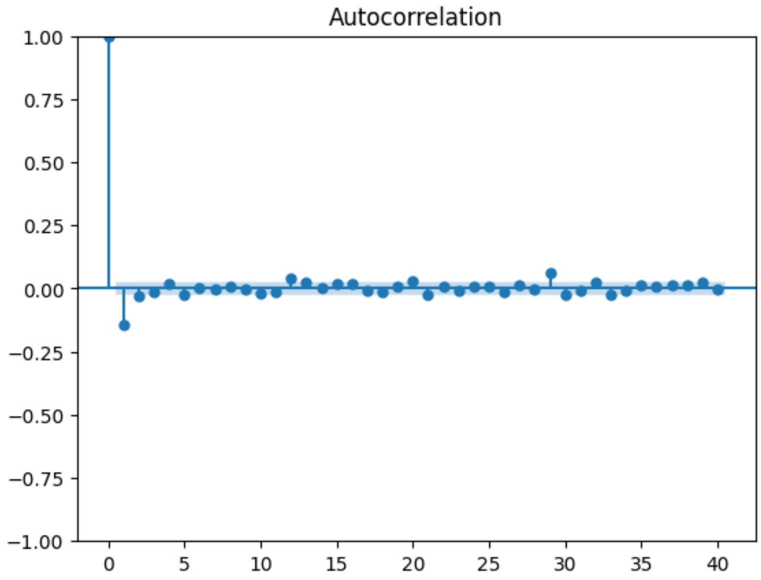

Figure 7.5: AutoCorrelation

Above diagram 7.5 shows the autocorrelation of a time series is displayed on the graph. The time series association with itself at various delays is known as the autocorrelation. The graphic indicates that the time series is not autocorrelated as the autocorrelation is very low for all delays. This indicates that the time series's future values are not significantly influenced by its past values. The graphic illustrates the data's absence of autocorrelation visually.

CHAPTER 8

Results and Discussion

The given data provide light on how well various forecasting models perform in predicting the price of crude oil. In order to determine if the time series data is stationary, the analysis first applies the Augmented Dickey-Fuller (ADF) test. Then, using a variety of performance criteria, the forecasting accuracy of three models is assessed: the Gated Recurrent Unit (GRU), the Convolutional Neural Network (CNN), and the AutoRegressive Integrated Moving Average (ARIMA). The results demonstrate how well these models can forecast crude oil prices; GRU performs best in terms of accuracy and consistency, while CNN and ARIMA show mixed results.

Parameter	value
ADF	-2.75
p value	0.06
critical value(1 percent)	-3.43
critical value(5 percent)	-2.86
critical value(10 percent)	-2.56

Table 8.1: AD Fuller data is non stationary for Crude Oil Price Forecasting

Table 8.1 shows a statistical test called the Augmented Dickey-Fuller (ADF) test is used to assess if a time series is stationary, or if its statistical characteristics do not vary over time. Here, the ADF test results for Crude Oil Price Forecasting indicate a p-value of 0.06 and

an ADF statistic of -2.75. At various degrees of significance, the test's critical values are as follows: -3.43 at 1 percent, -2.86 at 5 percent, -2.56 at 10 percent
Compare the ADF statistic with the essential levels to understand these findings. The crucial value at the 5 percent significance level is -2.86. As the ADF statistic (-2.75) exceeds the crucial threshold, the null hypothesis of non-stationarity is not rejected at this stage. Furthermore, the null hypothesis cannot be rejected at the 5 percent level since the p-value of 0.06 is higher than the 0.05 cutoff.
Consequently, the data indicates that the crude oil price time series is non-stationary, which means that its statistical characteristics, such mean and variance, fluctuate with time. Before using the majority of forecasting models, this non-stationarity needs to be corrected, frequently by differencing or other changes.

Parameter	value
ADF	-12.75
p value	8.2041e-24
critical value(1 percent)	-3.43
critical value(5 percent)	-2.86
critical value(10 percent)	-2.56

Table 8.2: AD Fuller data is stationary for Crude Oil Price Forecasting

Table 8.2 The findings of the Crude Oil Price Forecasting Augmented Dickey-Fuller (ADF) test indicate an ADF statistic of -12.75 and a p-value of 8.2041e-24.
We evaluate these findings by contrasting the ADF statistic with the crucial values. All of the crucial values, including the 1 percent threshold (-3.43), are substantially higher than the ADF statistic (-12.75). This indicates that we have a high degree of confidence in rejecting the non-stationarity null hypothesis.
Additionally, there is substantial evidence to refute the null hypothesis of non-stationarity, since the p-value (8.2041e-24) is very low and far lower than the usual significance values (0.01, 0.05, and 0.10). As a result, the findings imply that the time series for the price of crude oil is stable, meaning that its statistical characteristics, including mean and variance, do not alter over time. For many time series forecasting models, this stationarity is essential because

it guarantees that the connections in the data will continue to be stable across time, producing forecasts that are more accurate and dependable.

Method	RMSE	MSE	MAPE	R2 Square
GRU Model	9.76	95.40	6.99	0.49
CNN	3.006	9.03	6.46	0.91
ARIMA	14.06	197.76	25.23	-0.05

Table 8.3: Performance evaluation for Model at Validation Data in Crude Oil Price Forecasting

Table 8.3 show the three models used to anticipate crude oil prices are GRU (Gated Recurrent Unit), CNN (Convolutional Neural Network), and ARIMA (AutoRegressive Integrated Moving Average). The performance assessment metrics of each model are shown in the table. Among the measurements are R^2 (R-Squared), MAPE (Mean Absolute Percentage Error), MSE (Mean Squared Error), and RMSE (Root Mean Squared Error).
With the lowest RMSE (3.006) and MSE (9.03), the CNN model performs better than the other models, suggesting improved accuracy and reduced prediction errors. With a MAPE = 6.46, it indicates that its forecasts typically differ by 6.46 percent from the observed values. Additionally, the high R2 value of 0.91 shows that 91 percent of the variation in crude oil prices can be explained by the CNN model, demonstrating its potent predictive power.
Compared to the CNN model, the GRU model performs somewhat better but is less accurate, with an RMSE of 9.76 and an MSE of 95.40. Its R2 value of 0.49 suggests moderate explanatory power, while its MAPE of 6.99 indicates a somewhat greater average percentage error.
With the greatest RMSE (14.06) and MSE (197.76), which indicate significant prediction errors, the ARIMA model performs the poorest. Significant departures from the predictions are highlighted by its MAPE of 25.23, and the negative R2 value (-0.05) indicates that the predictive power was inadequate and failed to identify the underlying trends in the data.

Method	RMSE	MSE	MAPE	R2 Square
GRU Model	3.99	15.95	13.75	0.497
CNN	4.12	16.98	14.86	0.84
ARIMA	14.91	222.35	11.05	-0.57

Table 8.4: Performance evaluation for Model at Testing Data in Crude Oil Price Forecasting

Table 8.4 shows the GRU (Gated Recurrent Unit), CNN (Convolutional Neural Network), and ARIMA (AutoRegressive Integrated Moving Average) models' performance evaluations on testing data for crude oil price predictions are shown in the table. Among the measurements are R^2 (R-Squared), MAPE (Mean Absolute Percentage Error), MSE (Mean Squared Error), and RMSE (Root Mean Squared Error).

With the lowest RMSE (3.99) and MSE (15.95), which show it has the least average prediction error and variance, the GRU model performs the best out of all the models. With a reasonable accuracy in percentage terms shown by its MAPE of 13.75 and a R^2 value of 0.497, this model accounts for over half of the variation in crude oil prices.

The CNN model likewise performs rather well, despite having somewhat higher MSE (16.98) and RMSE (4.12). Its MAPE of 14.86, however, is more than the GRU model's, suggesting a greater average percentage error. Notwithstanding its higher RMSE and MAPE values, the CNN model exhibits significant explanatory power with a R^2 value of 0.84, explaining 84 percent of the variance.

With an MSE of 222.35 and an RMSE of 14.91, the ARIMA model performs the poorest, indicating significant prediction errors. It's interesting to note that, although having the lowest MAPE of 11.05, which suggests superior performance in percentage terms, its negative R^2 value of -0.57 indicates that it performs worse than a simple mean model in terms of capturing the data patterns.

Date	GRU	CNN	ARIMA	Real World Data
24-07-16	79.27	80.44	81.62	80.76
24-07-17	79.57	80.85	80.83	82.84
24-07-18	79.64	81.27	80.90	82.82
24-07-19	79.60	81.49	81.73	80.12
24-07-20	79.68	81.65	81.54	80.28
24-07-21	79.79	81.81	83.37	79.82
24-07-22	80.21	82.13	82.80	79.77
24-07-23	80.60	82.25	83.87	76.95
24-07-24	81.08	82.58	83.16	77.58
24-07-25	81.38	82.77	82.33	78.27
24-07-26	81.37	83.23	81.41	77.16
24-07-27	81.06	83.65	82.09	77.34
24-07-28	80.80	83.30	82.62	76.83
24-07-29	80.73	83.65	82.20	75.80
24-07-30	80.71	83.79	81.91	74.73

Table 8.5: Future Forecasting Crude Oil Price for next 14 days

The table 8.5 shows when the forecasting models:GRU, CNN, and ARIMA are compared to actual crude oil price data, different forecasting models perform differently during the given time period.

- GRU (Gated Recurrent Unit): The GRU model closely tracks changes in actual prices and shows a somewhat consistent forecasting ability. For example, on July 24 to 26, it forecasts a price of 81.37, which is quite near to the 77.16 actual price. This suggests that the temporal connections in the data are adequately captured by GRU.
- Convolutional Neural Network, or CNN: This model predicts outcomes marginally

more accurately than ARIMA but less accurately than GRU. It predicts values that are somewhat in line with actual prices, indicating that it recognizes patterns well but could have trouble with certain data volatility.

- ARIMA (AutoRegressive Integrated Moving Average): When market circumstances are turbulent, the ARIMA model tends to lag behind real prices. For instance, its prediction for July 24th, 2023, at 83.87, is higher than the actual value of 76.95, suggesting that ARIMA could not adjust to sudden changes in the market swiftly.

All things considered, the GRU model seems to offer the most accurate and consistent predictions, which makes it a better option for time series forecasting of crude oil prices, particularly in settings when the market is dynamic. CNN does a great job, while ARIMA has trouble catching sudden changes in the market.

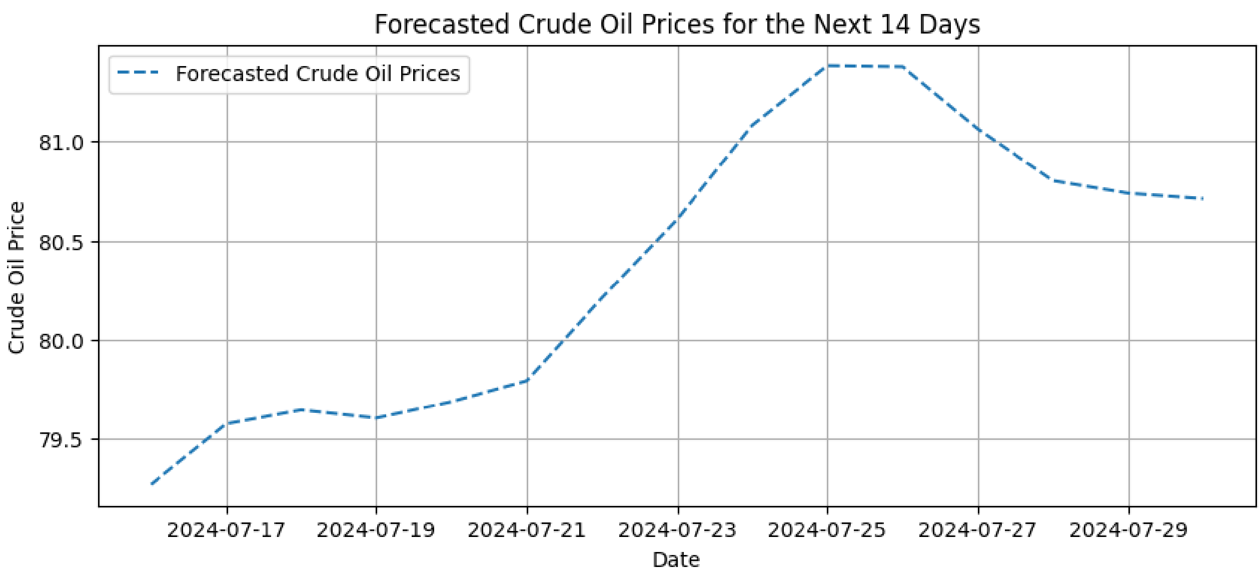

Figure 8.1: Crude Oil Price 15 days forecasting value using GRU Model

Above diagram 8.1 shows the predicted price of crude oil for the next 14 days is displayed on the graph. It is anticipated that the rates would increase from 79.25 dollars to a maximum of 81.75 dollars on July 25 and then drop to 80.75 dollars on July 29. This points to a possible time frame in which oil prices may rise before somewhat declining

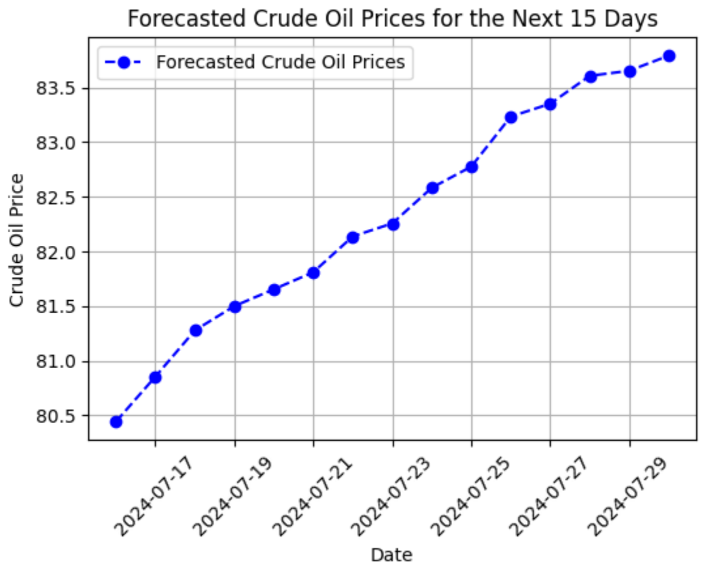

Figure 8.2: Crude Oil Price 15 days forecasting value using CNN Model

Above diagram 8.2 shows the graph displays the anticipated rise in crude oil prices for the ensuing fifteen days. Starting at 80.50 dollars, the price increases gradually to 83.80 dollars at the conclusion of the projected period. This points to a promising trend in the oil market, where price increases are anticipated to continue

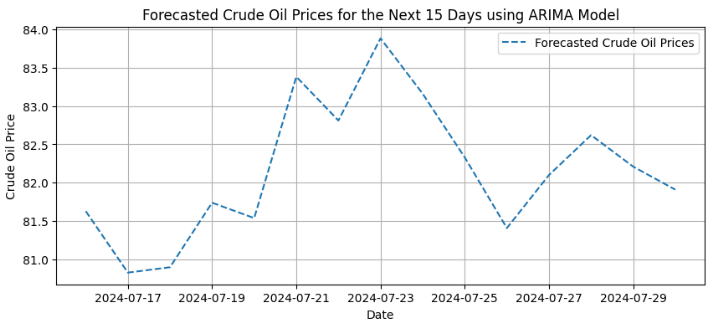

Figure 8.3: Crude Oil Price 15 days forecasting value using ARIMA Model

Above diagram 8.3 shows the graph which uses an ARIMA model to predict the price of crude oil for the following fifteen days. The price is predicted to vary over the following two weeks, with a sharp increase around July 23rd and a subsequent decrease. After that, it is anticipated that the price will rise once more, peaking on July 27th, and then begin to decline as the forecast period draws to an end

8.1 Global Events Effect on Crude Oil price

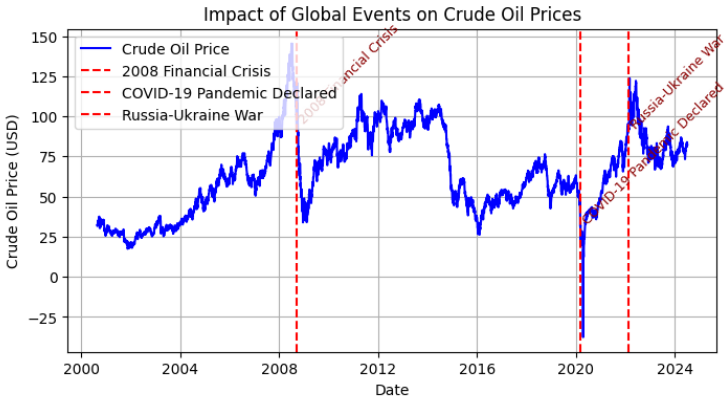

Figure 8.4: Crude Oil Price Effect during Global Events

Above diagram 8.4 shows major world events explain much on the crude oil price in 2000 to 2024. It is from the graph, that economic as well as the geopolitical arena discontinuity brings great changes in price crude oil.

Oil prices collapsed during the 2008 Financial Crisis from highs above 140 dollars per barrel, as it contracted the global economy and cut down on demand for energy, thereby forcing prices to spiral down.

The COVID-19 pandemic led to an unprecedented collapse in crude oil prices in 2020. Demand crashed due to lockdowns worldwide, stopped travel, and reduced industrial activities. For the first time in history, oil prices turned negative in some markets due to oversupply and storage limitations.

Oil price spikes to extreme highs in 2022 during the Russia-Ukraine War, where this disrupted the supply of energy supplies in the continent, primarily Europe, heavily relying on Russia for oil and gas supplies. Such fear of scarcity, combined with international sanctions on Russia, has priced oil at 120 per dollars barrel and more.

These events are the epitome of the volatile nature of crude oil prices, which are molded by global economic health, dynamics of demand-supply, and geopolitical tensions. Oil remains a critical global commodity, and its price is intricately linked to the major global events.

8.2 Comparison between neural network and statistical approaches

From the results, the neural network models, such as GRU and CNN, outperform the statistical approach of ARIMA in the forecast of crude oil prices. The GRU model has shown the best accuracy and consistency with a lower RMSE of 3.99 on testing data and robust in predicting temporal dependencies. CNN also did very well, with a high R2 value of 0.84, showing its ability in capturing patterns, though slightly higher errors compared to GRU.
On the other hand, ARIMA exerts the poorest result with the highest RMSE of 14.91 on testing data and a negative R2 value of -0.57, hence unable to grasp nonlinear and dynamic trends of crude oil prices. Although ARIMA could work fine with stationary data, it failed to work well on the volatile nature of this market.
Overall, GRU and CNN tend to adapt more to complex time series patterns; therefore, both are reliable in the prediction of crude oil price, while ARIMA's linear approach to time series modeling is limited for this scenario.

CHAPTER 9

Conclusion

Studies on the prediction of crude oil prices with CNN (Convolutional Neural Network), ARIMA (AutoRegressive Integrated Moving Average), and GRU (Gated Recurrent Unit) advanced machine learning models provide important new information about how well these approaches work for time series forecasting.

How to optimize hyperparameters to improve the performance of deep learning models was one of the main research problems that was investigated. The results suggest that, even though the GRU and CNN models have demonstrated encouraging outcomes when using their default hyperparameters, more tweaking may improve their performance metrics. In comparison to ARIMA's RMSE of 14.91, the models testing stage Root Mean Square Error (RMSE) values of 3.99 (GRU) and 4.12 (CNN) demonstrate a significant decrease in error.Techniques for hyperparameter tweaking In order to increase accuracy and predictive power, grid search is used to methodically identify the ideal values for learning rates, batch sizes, and layer configurations.

Comparison of Conventional and Deep Learning Models: The second study topic assessed the robustness and accuracy of the conventional ARIMA model vs the deep learning models (CNN and GRU). The findings show that CNN outperformed ARIMA, with an R-squared of -0.57 and a Mean Absolute Percentage Error (MAPE) of 14.86 during testing, indicating a definite advantage for CNN. While ARIMA suffered because of its linear assumptions, the GRU model also showed competitive performance, suggesting that deep learning approaches are more competent at capturing the nonlinear patterns present in crude oil price data.This means that deep learning models which perform better in volatile situations like the crude oil markets are better suited for complicated time series forecasting.

Although the primary emphasis of the current study was historical pricing data, future research ought to look at adding these external factors to the models as extra features. The explanatory power of the models, for example, may be greatly increased by include data on economic indices like GDP growth rates, OPEC actions, and geopolitical concerns. In-depth research may be possible to determine which factors most significantly affect crude oil prices, which might result in forecasting that is even more precise.

The results show that the precision of the CNN and GRU models can offer insightful information to investors, decision-makers, and oil-dependent companies. Precise forecasts of prices help mitigate the risks associated with price volatility by providing valuable insights for strategic choices including budget allocation, investment scheduling, and inventory management. Stakeholders could use these models, for example, to protect themselves against price decreases or to maximize their buying tactics when prices are expected to rise. These predictive models' operational use may improve decision-making procedures, which would eventually improve financial results.

This study demonstrates how well deep learning models in particular, CNN and GRU predict crude oil prices. In order to increase these models accuracy and resilience, more research should be done on hyperparameter tuning, since it can lead to large performance gains. Applying systematic tuning techniques, such Bayesian optimization or grid search, might reveal ideal setups that improve prediction performance.

Furthermore, it is imperative that external variables like market mood, economic statistics, and geopolitical developments be included into the models.This integration might lead to enhanced accuracy by providing a more thorough understanding of the factors impacting crude oil pricing. To assess their effectiveness versus GRU and CNN, comparison studies using more sophisticated forecasting techniques, such as LSTM networks or hybrid models that include conventional and deep learning methods, should be conducted.

Investigating the use of real time data streams can facilitate the creation of dynamically updating models that can produce forecasts that are more sensitive to changes in the market. To evaluate these models' scalability and incorporation into decision-support systems for oil market stakeholders, it is imperative to examine their implementation in real world scenarios. Events like wars, pandemics, and financial crises affect the crude oil price due to a supply and-demand situation. For instance, wars disrupt production and transportation, hence a shortage in supply and leading to increased prices, such as in the war between Russia and Ukraine. Pandemics like COVID-19 reduce global economic activity and hence sharply lower oil demand and prices, as was seen in 2020. Financial crises, like the 2008 global meltdown,lower industrial output and, consequently, energy consumption, driving prices lower. Events such as these introduce volatility and uncertainty in the market, and policymakers and traders have to be keenly aware of their impact on crude oil dynamics as part of an effective decision-making process.

Chapter 10

Appendix

Github Link

Bibliography

[Abdollahi and Ebrahimi, 2020] Abdollahi, H. and Ebrahimi, S. B. (2020). A new hybrid model for forecasting brent crude oil price. *Energy*, 200:117520.

[Adebiyi et al., 2014] Adebiyi, A. A., Adewumi, A. O., and Ayo, C. K. (2014). Comparison of arima and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*, 2014(1):614342.

[Aghdam et al., 2017] Aghdam, H. H., Heravi, E. J., et al. (2017). Guide to convolutional neural networks. *New York, NY: Springer*, 10(978-973):51.

[Amoco, 2001] Amoco, B. (2001). *BP Statistical Review of World Energy, June 2001*. BP Amoco.

[Balas et al., 2020] Balas, V. E., Kumar, R., Srivastava, R., et al. (2020). *Recent trends and advances in artificial intelligence and internet of things*. Springer.

[Borovykh et al., 2017] Borovykh, A., Bohte, S., and Oosterlee, C. W. (2017). Conditional time series forecasting with convolutional neural networks. *arXiv preprint arXiv:1703.04691*.

[Chen et al., 2016] Chen, H., Liao, H., Tang, B.-J., and Wei, Y.-M. (2016). Impacts of opec's political risk on the international crude oil prices: An empirical analysis based on the svar models. *Energy Economics*, 57:42–49.

[Chen et al., 2018] Chen, Y., Zhang, C., He, K., and Zheng, A. (2018). Multi-step-ahead crude oil price forecasting using a hybrid grey wave model. *Physica A: Statistical Mechanics and its Applications*, 501:98–110.

[Cho et al., 2014] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

[Chung et al., 2014] Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

[Dees et al., 2007] Dees, S., Karadeloglou, P., Kaufmann, R. K., and Sanchez, M. (2007). Modelling the world oil market: Assessment of a quarterly econometric model. *Energy policy*, 35(1):178–191.

[Dey and Salem, 2017] Dey, R. and Salem, F. M. (2017). Gate-variants of gated recurrent unit (gru) neural networks. In *2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS)*, pages 1597–1600. IEEE.

[Elshendy et al., 2018] Elshendy, M., Colladon, A. F., Battistoni, E., and Gloor, P. A. (2018). Using four different online media sources to forecast the crude oil price. *Journal of Information Science*, 44(3):408–421.

[Few, 2013] Few, S. (2013). *Information Dashboard Design: Displaying data for at-a-glance monitoring*, volume 5. Analytics Press Burlingame, CA.

[García et al., 2015] García, S., Luengo, J., Herrera, F., et al. (2015). *Data preprocessing in data mining*, volume 72. Springer.

[Gholamian et al., 2005] Gholamian, M. R., Ghomi, S. F., and Ghazanfari, M. (2005). A hybrid systematic design for multiobjective market problems: a case study in crude oil markets. *Engineering Applications of Artificial Intelligence*, 18(4):495–509.

[Guidolin and Timmermann, 2009] Guidolin, M. and Timmermann, A. (2009). Forecasts of us short-term interest rates: A flexible forecast combination approach. *Journal of Econometrics*, 150(2):297–311.

[Hagen, 1994] Hagen, R. (1994). How is the international price of a particular crude determined? *OPEC Review*, 18(1):127–135.

[Hamilton, 2009] Hamilton, J. D. (2009). Causes and consequences of the oil shock of 2007-08. Technical report, National Bureau of Economic Research.

[He et al., 2010] He, A. W., Kwok, J. T., and Wan, A. T. (2010). An empirical model of daily highs and lows of west texas intermediate crude oil prices. *Energy Economics*, 32(6):1499–1506.

[Hochreiter and Schmidhuber, 1997] Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.

[Kim and Kim, 2019] Kim, T. and Kim, H. Y. (2019). Forecasting stock prices with a feature fusion lstm-cnn model using different representations of the same data. *PloS one*, 14(2):e0212320.

[Kulkarni and Haidar, 2009] Kulkarni, S. and Haidar, I. (2009). Forecasting model for crude oil price using artificial neural networks and commodity futures prices. *arXiv preprint arXiv:0906.4838*.

[Kumar et al., 2020] Kumar, Y. J. N., Preetham, P., Varma, P. K., Rohith, P., and Kumar, P. D. (2020). Crude oil price prediction using deep learning. In *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, pages 118–123. IEEE.

[Kurdi, 2022] Kurdi, G. (2022). Toward an electronic resource for systematic reviews in computer science.

[LeCun et al., 2015] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553):436–444.

[Li et al., 2024] Li, J., Hong, Z., Zhang, C., Wu, J., and Yu, C. (2024). A novel hybrid model for crude oil price forecasting based on meemd and mix-kelm. *Expert Systems with Applications*, 246:123104.

[Miao et al., 2017] Miao, H., Ramchander, S., Wang, T., and Yang, D. (2017). Influential factors in crude oil price forecasting. *Energy Economics*, 68:77–88.

[Plourde and Watkins, 1994] Plourde, A. and Watkins, G. (1994). How volatile are crude oil prices? *OPEC review*, 18(4):431–444.

[Ravanelli et al., 2018] Ravanelli, M., Brakel, P., Omologo, M., and Bengio, Y. (2018). Light gated recurrent units for speech recognition. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2(2):92–102.

[Sadorsky, 2006] Sadorsky, P. (2006). Modeling and forecasting petroleum futures volatility. *Energy economics*, 28(4):467–488.

[Safari and Davallou, 2018] Safari, A. and Davallou, M. (2018). Oil price forecasting using a hybrid model. *Energy*, 148:49–58.

[Samek et al., 2021] Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., and Müller, K.-R. (2021). Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE*, 109(3):247–278.

[Shambulingappa et al., 2020] Shambulingappa, H. et al. (2020). Crude oil price forecasting using machine learning. *International Journal of Advanced Scientific Innovation*, 1(1):1–11.

[Stock and Watson, 2006] Stock, J. H. and Watson, M. W. (2006). Forecasting with many predictors. *Handbook of economic forecasting*, 1:515–554.

[Tang et al., 2015] Tang, L., Dai, W., Yu, L., and Wang, S. (2015). A novel ceemd-based eelm ensemble learning paradigm for crude oil price forecasting. *International Journal of Information Technology & Decision Making*, 14(01):141–169.

[Tsantekidis et al., 2017] Tsantekidis, A., Passalis, N., Tefas, A., Kanniainen, J., Gabbouj, M., and Iosifidis, A. (2017). Forecasting stock prices from the limit order book using convolutional neural networks. In *2017 IEEE 19th conference on business informatics (CBI)*, volume 1, pages 7–12. IEEE.

[Valueva et al., 2020] Valueva, M. V., Nagornov, N., Lyakhov, P. A., Valuev, G. V., and Chervyakov, N. I. (2020). Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. *Mathematics and computers in simulation*, 177:232–243.

[Venkatesan and Li, 2017] Venkatesan, R. and Li, B. (2017). *Convolutional neural networks in visual computing: a concise guide*. CRC Press.

[Wan and Kao, 2015] Wan, J.-Y. and Kao, C.-W. (2015). Interactions between oil and financial marketsâdo conditions of financial stress matter? *Energy Economics*, 52:160–175.

[Wang et al., 2018] Wang, M., Tian, L., and Zhou, P. (2018). A novel approach for oil price forecasting based on data fluctuation network. *Energy Economics*, 71:201–212.

[Wex et al., 2013] Wex, F., Widder, N., Liebmann, M., and Neumann, D. (2013). Early warning of impending oil crises using the predictive power of online news stories. In *2013 46th hawaii international conference on system sciences*, pages 1512–1521. IEEE.

[Witten et al., 2005] Witten, I. H., Frank, E., Hall, M. A., Pal, C. J., and Data, M. (2005). Practical machine learning tools and techniques. In *Data mining*, volume 2, pages 403–413. Elsevier Amsterdam, The Netherlands.

[Xiong et al., 2013] Xiong, T., Bao, Y., and Hu, Z. (2013). Beyond one-step-ahead forecasting: evaluation of alternative multi-step-ahead forecasting models for crude oil prices. *Energy Economics*, 40:405–415.

[Yahoo Inc., 2023] Yahoo Inc. (2023). Yahoo finance api. Accessed: 30/06/2024.

[Yu et al., 2016] Yu, L., Dai, W., and Tang, L. (2016). A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting. *Engineering Applications of Artificial Intelligence*, 47:110–121.

[Yu et al., 2008] Yu, L., Wang, S., and Lai, K. K. (2008). Forecasting crude oil price with an emd-based neural network ensemble learning paradigm. *Energy economics*, 30(5):2623–2635.

[Yu et al., 2015] Yu, L., Wang, Z., and Tang, L. (2015). A decomposition–ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting. *Applied Energy*, 156:251–267.

[Zhang et al., 1998] Zhang, G., Patuwo, B. E., and Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting*, 14(1):35–62.

[Zhang et al., 2015] Zhang, J.-L., Zhang, Y.-J., and Zhang, L. (2015). A novel hybrid method for crude oil price forecasting. *Energy Economics*, 49:649–659.