# AN INTELLIGENT DECISION SUPPORT SYSTEM FOR CRUDE OIL TRADING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS OF
THE MIDDLE EAST TECHNICAL UNIVERSITY
BY

BURCU KOÇ GÖLTAŞ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF INFORMATION SYSTEMS

SEPTEMBER 2023

Approval of the thesis:

**AN INTELLIGENT DECISION SUPPORT SYSTEM FOR CRUDE OIL TRADING**

Submitted by BURCU KOÇ GÖLTAŞ in partial fulfillment of the requirements for the degree of **Master of Science in Information Systems Department, Middle East Technical University** by,

Prof. Dr. Banu Günel Kılıç
Dean, **Graduate School of Informatics** ____________________

Prof. Dr. Altan Koçyiğit
Head of Department, **Information Systems Dept., METU** ____________________

Assoc. Prof. Dr. P. Erhan Eren
Supervisor, **Information Systems Dept., METU** ____________________

**Examining Committee Members:**

Prof. Dr. Altan Koçyiğit
**Information Systems Dept., METU** ____________________

Assoc. Prof. Dr. P. Erhan Eren
**Information Systems Dept., METU** ____________________

Asst. Prof. Dr. Serhat Peker
**Management Information Systems Dept., İzmir Bakırçay University** ____________________

**Date:** *06.09.2023*

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

**Name, Last name : Burcu Koç Göltaş**

**Signature :**

## ABSTRACT

**AN INTELLIGENT DECISION SUPPORT SYSTEM FOR CRUDE OIL TRADING**

Koç Göltaş, Burcu

MSc., Department of Information Systems

Supervisor: Assoc. Prof. Dr. P. Erhan Eren

September 2023, 106 pages

Crude oil is one of the most traded commodities in the world. Traders engaged in the crude oil market for various purposes, such as generating profits and hedging price-related risks. Crude oil price is highly volatile as it is affected by a wide range of factors such as economic factors, political events, prices of other commodities, and financial instruments. Therefore, the biggest obstacle to profitable trading is the difficulty of keeping track of all factors affecting the market and the risks associated with price volatility. As a solution, this study examines all the factors that affect the crude oil market and provides an architecture for aggregating and processing these factors. It also holistically integrates technical, fundamental, and sentiment analysis methodologies and provides intelligent methods for their application to support crude oil trading decisions. A prototype of the proposed system was also implemented to demonstrate its applicability. This prototype demonstrates that predictive models accurately predict crude oil prices and that even the most basic trading methods may produce profitable outcomes.

Keywords: Crude Oil, Financial Decision Support Systems, Fundamental Analysis, Technical Analysis, Design Science Research

# ÖZ

## HAM PETROL TİCARETİ İÇİN AKILLI KARAR DESTEK SİSTEMİ

Koç Göltaş, Burcu

Yüksek Lisans, Bilişim Sistemleri Bölümü

Tez Yöneticisi: Doç. Dr. P. Erhan Eren

Eylül 2023, 106 sayfa

Ham petrol, dünyada en çok ticareti yapılan emtialardan biridir. Yatırımcılar, kâr elde etmek ve fiyatla ilgili risklerden korunmak gibi çeşitli amaçlarla ham petrol piyasasında işlem yapmaktadır. Ham petrol fiyatı, ekonomik faktörler, siyasi olaylar, diğer emtiaların ve finansal araçların fiyatları gibi çok çeşitli faktörlerden etkilendiği için oldukça değişkendir. Bu nedenle, kârlı yatırımların önündeki en büyük engel, piyasayı etkileyen tüm faktörleri ve fiyat oynaklığı ile ilişkili riskleri takip etmenin zorluğudur. Çözüm olarak, bu çalışma ham petrol piyasasını etkileyen tüm faktörleri incelemekte ve bu faktörleri bir araya getirip işlemek için bir mimari sunmaktadır. Ayrıca teknik analiz, temel analiz ve duyarlılık analizi metodolojilerini bütünsel bir şekilde entegre etmekte ve ham petrol alım satım kararlarını desteklemek için bu analizlerin uygulanmasına yönelik akıllı yöntemler sunmaktadır. Önerilen sistemin uygulanabilirliğini göstermek için bir prototipi de hayata geçirilmiştir. Bu prototip, tahmin modellerinin ham petrol fiyatlarını doğru bir şekilde tahmin ettiğini ve en temel yatırım yöntemlerinin bile kârlı sonuçlar üretebileceğini göstermektedir.

Anahtar Sözcükler: Ham Petrol, Finansal Karar Destek Sistemleri, Temel Analiz, Teknik Analiz, Bilim-Tasarım Yönetimi

To My Family

## ACKNOWLEDGMENTS

First of all, I would like to thank my supervisor, Assoc. Prof. Dr. P. Erhan Eren for his guidance and valuable contributions throughout this thesis. I am also grateful for his valuable collaborations that helped me get introduced to this field and find the opportunity to improve myself.

Besides my supervisor, I would like to thank my colleagues, especially Alp Bayar, Arif Ozan Kızıldağ, and Kerem Nazlıel, whose doors I could come knocking on without hesitation during my thesis process for their help, encouragement, and friendship. I would also like to thank all the administrative staff of our department, especially Ayşenur Özdere Yüksel and Sibel Gülnar, for their endless support.

I would also like to thank my dear friends Cansu Güney, Dilara Keser, Ela Sümeyye Seçim, Elif Bayat, Faruk Sağlamöz, Melis Çeviker, Mizgin Akın, and Tunç Siper for being my companions and for the encouragement they gave me in this challenging journey.

I would like to thank my dearest mother, father, and siblings for always being there for me, trying for my happiness, and picking me up when I was down. I'm so lucky to have you.

Last but not least, I want to thank the love of my life and best friend, Yılmaz Taylan Göltaş, for accompanying me on this journey and never letting go of my hand. This research would not have been possible without him. I am very thrilled about our future adventures and the life we will share.

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| **ABSA** | Aspect-Based Sentiment Analysis |
| **ADX** | Average Directional Index |
| **AI** | Artificial Intelligence |
| **ANFIS** | Adaptive Neuro-Fuzzy Inference System |
| **ARIMA** | Autoregressive Integrated Moving Average |
| **ARMA** | Autoregressive Moving Average |
| **ATR** | Average True Range |
| **BERT** | Bidirectional Encoder Representations from Transformer |
| **BiGRU** | Bidirectional Gated Recurrent Unit |
| **BMA** | Bayesian Model Averaging |
| **BOW** | Bag of Word |
| **BPNN** | Backpropagation Neural Network |
| **CCI** | Commodity Channel Index |
| **CEEMDAN** | Complete Ensemble Empirical Mode Decomposition with Adaptive Noise |
| **CME** | Chicago Mercantile Exchange |
| **DMKL** | Deep Multiple Kernel Learning |
| **DNN** | Deep Neural Network |
| **DSR** | Design Science Research |
| **DSRM** | Design Science Research Methodology |
| **DSS** | Decision Support System |
| **EEMD** | Ensemble Empirical Mode Decomposition |
| **EIA** | Energy Information Administration |
| **ELM** | Extreme Learning Machine |
| **EMA** | Exponential Moving Average |
| **EMD** | Empirical Mode Decomposition |
| **EPU** | Economic Policy Uncertainty |
| **FLANN** | Functional Link Artificial Neural Network |
| **FRED** | Federal Reserve Economic Data |
| **GARCH** | Generalized Autoregressive Conditional Variance |
| **GBDT** | Gradient Boosting Decision Tree |
| **GPT** | Generative Pre-Trained Transformer |
| **GRU** | Gated Recurrent Unit |
| **GSCI** | Goldman Sachs Commodity Index |
| **HDFS** | Hadoop Distributed File System |

| | |
|---|---|
| **ICE** | Intercontinental Exchange Group |
| **IISEC** | International Informatics and Software Engineering Conference |
| **IMF** | Intrinsic Mode Function |
| **IS** | Information Systems |
| **KELM** | Kernel Extreme Learning Machine |
| **KNN** | k-Nearest Neighbors |
| **KPCA** | Kernel Principal Component Analysis |
| **LASSO** | Least Absolute Shrinkage and Selection Operator |
| **LDA** | Latent Dirichlet Allocation |
| **LGBM** | Light Gradient Boosting Machine |
| **LLE** | Locally Linear Embedding |
| **LLM** | Large Language Model |
| **LR** | Linear Regression |
| **LSTM** | Long Short-Term Memory |
| **MA** | Moving Average |
| **MACD** | Moving Average Convergence-Divergence |
| **MAE** | Mean Absolute Error |
| **MAST** | Multifaceted Analysis of Subjective Text |
| **MEMD** | Multivariate Empirical Mode Decomposition |
| **MLR** | Multiple Linear Regression |
| **MSCI** | Morgan Stanley Capital International |
| **MSE** | Mean Squared Error |
| **MWRNN** | Multiple Wavelet Recurrent Neural Network |
| **NLP** | Natural Language Processing |
| **NYMEX** | New York Mercantile Exchange |
| **OBV** | On-Balance Volume |
| **ODEE** | Open Domain Event Extraction |
| **OPEC** | Organization of the Petroleum Exporting Countries |
| **PSO** | Particle Swarm Optimization |
| **RBSA** | Rule-based Sentiment Analysis |
| **ReLU** | Rectified Linear Unit |
| **RF** | Random Forest |
| **RL** | Reinforcement Learning |
| **RNN** | Recurrent Neural Network |
| **ROC** | Rate of Change |
| **RSI** | Relative Strength Index |

| | |
|---|---|
| **RVFL** | Random Factor Functional Link Network |
| **S&P** | Standard & Poor |
| **SDAE** | Stacked Denoising Autoencoders |
| **SMA** | Simple Moving Average |
| **SOA** | Seagull Optimization Algorithm |
| **SPR** | Strategic Petroleum Reserve |
| **SVM** | Support Vector Machine |
| **TF-IDF** | Term Frequency - Inverse Document Frequency |
| **USO** | United States Oil |
| **VADER** | Valence Aware Dictionary and Sentiment Reasoner |
| **VAR** | Vector Autoregression |
| **VECM** | Vector Error Correction Model |
| **VIX** | Volatility Index |
| **VMD** | Variational Mode Decomposition |
| **WNN** | Wavelet Neural Network |
| **WOA** | Whale Optimization Algorithm |
| **WTI** | West Texas Intermediate |
| **XGBoost** | Extreme Gradient Boosting |
| **XOM** | Exxon Mobile Corporation |

# CHAPTER 1

# INTRODUCTION

## 1.1 Research Background

Crude oil is one of the most popular commodities worldwide; therefore, fluctuations in its price are closely monitored globally [1]. Crude oil prices depend on various parameters, including global economic conditions, the balance between supply and demand, and geopolitical events [2]. As a result, the crude oil price tends to fluctuate over time. Governments, organizations, and individuals trade in the crude oil market to hedge against and profit from this price volatility. A broad range of entities participate in the oil market by utilizing various instruments in spot and derivatives markets. Individual traders often trade in derivatives markets because derivatives enable them to generate profits without physical delivery [3].

The world has approximately 160 distinct crude oil grades, many traded on various marketplaces [4]. WTI (West Texas Intermediate) and Brent (Brent) are the most prominent of these. Due to the vast diversity of crude oils and the fact that many variables influence their price, traders' most significant challenge is the volatility of price movements and the constant need to analyze data from a broad range of sources to track the fluctuations. For this reason, creating models that accurately predict price variations to some extent by considering the factors affecting the crude oil price is of great importance for traders and increases profitability [5].

There are several ways to measure financial product prices and profitability. The first is fundamental analysis. This is the process of predicting the price of a financial instrument by considering the equilibrium of supply and demand, as well as other economic factors that can influence the price [6]. On the other hand, technical analysis is the process of analyzing a financial instrument's historical price using various techniques and attempting to identify patterns to gain insight into current and future price movements [7]. In addition, a broad range of techniques are employed to assess investor sentiment, which is utilized to gain insight into market trends [8]. The literature lacks studies that provide a perspective that includes all of these analyses for trading decisions of a financial instrument, and this study aims to fill this gap.

## 1.2 Research Aim and Objectives

The main objective of this study is to propose a decision support system for individual traders who trade in the crude oil market for profit and who cannot fully utilize the data in the market. The primary advantages of the system are the models that forecast the price of various trading instruments, the specification of the data types to be fed into these models, and the development of an infrastructure for collecting and aggregating this broad range of data. The system provides price forecasts and incorporates technical analysis techniques, back testing for technical trading rules, and the impact of news flows on crude oil prices can also be observed. The research objectives of the study were defined based on these objectives:

- To specify the data requirements of the decision support system for crude oil trading by utilizing a systematic literature review.

- To propose a framework to support decisions on crude oil trading.

- To develop the implementation of the proposed framework to support crude oil trading decisions.

## 1.3 Significance of the Study

This study made several contributions to the literature, which are presented below.

- All available data sources related to crude oil trading were compiled through a systematic literature review.

- A decision support system for collecting and preprocessing various forms of data from various sources has been proposed, and this system brings together many types of analysis utilized in crude oil trading from a holistic perspective.

- A prototype of the proposed system has been constructed, and the applicability and functionality of the available data types and models has been validated.

## 1.4 Structure of the Thesis

This study consists of six chapters in which the proposed system is detailed by considering the research objectives determined to solve the problems stated in this section. After the introductory chapter, the literature review describes the different instruments in the crude oil market, the market participants, the factors affecting the price, and the types of analyses that can be used to understand the market. After that, the decision support system literature is examined in detail, and the decision support systems used in the finance domain and the systems designed to be used in crude oil trading are analyzed. Chapter 3 sets out the study's research questions before delving into the research methodology employed to address them. Chapter 4 demonstrates the proposed system's details, such as its architecture, data types, and models that can be employed. Chapter 5 provides an overview of the prototype of the proposed system. The final chapter, Chapter 6, presents the results of the study, its limitations, contributions, and future works.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Crude Oil Trading

One of the most widely traded commodities worldwide is crude oil [1]. Over 100 million barrels of crude oil are traded daily on the exchanges [9]. Crude oil trading considerably impacts the global economy and financial markets [10]. This is because crude oil price volatility significantly impacts international politics and economies, directly affecting people's earnings and the growth of entire nations [2]. The remaining sections of this chapter cover the exchanges, actors, and factors influencing the market in crude oil trade.

### 2.1.1 New York Mercantile Exchange

Crude oil futures are traded on a variety of exchanges. NYMEX is one of the major exchanges for commodities futures owned by Chicago-based CME. NYMEX's predecessor, the New York Butter and Cheese Exchange, was founded in 1882 and renamed the New York Mercantile Exchange in 1888 [11]. Oil, natural gas, metals, and agricultural products are among the many commodities traded on the NYMEX through futures contracts [11]. Commodities prices traded on the NYMEX serve as a benchmark for international markets due to the exchange's prominence in the world's commodities markets. WTI, one of the most important crude oils, is traded on this exchange.

### 2.1.2 Chicago Mercantile Exchange

It is a global derivatives exchange that enables trading in various financial and commodity products. It was founded as the Chicago Butter and Egg Board on December 17, 1898, and was renamed Chicago Mercantile Exchange in 1919 [12]. Agriculture-related commodities were exchanged mostly at first, and there are many different futures contracts available today, such as Bitcoin futures [13], EUR/USD futures [14], and maize and soybean futures [15].

### 2.1.3 Intercontinental Exchange Group

With its headquarters in the United States, the Intercontinental Exchange Group (ICE) is a multinational company that offers services for various asset classes, including commodities, energy, equities, interest rates, and currencies [16]. The ICE was established on May 11, 2000, and while its initial focus was on electronic energy trading, it now provides various products. Future exchanges have been established separately for the US, Europe, Singapore, and Abu Dhabi [10]. Brent, one of the most important crude oils, is traded on ICE Futures Europe.

### 2.1.4 Type of Traders

Trading in crude oil involves many market participants, such as speculators, oil-importing nations, oil firms, individual refineries, and oil-producing nations [17]. These players might meet their needs for physical oil, manage price risk, and profit from speculation on price changes. For instance, oil companies seek to sell their product for the highest possible price. Speculators, hedgers, and arbitrageurs are the three major types of traders in the market.

*Speculators:* One of the key players in the crude oil market is speculators. Without purchasing physical products, they attempt to profit by considering market price movements [18]. They make decisions based on market fundamentals, technical indicators, geopolitical factors, and market sentiment [19]. They may take long or short positions according to whether commodity prices are expected to rise or fall [18]. Speculators contribute to price discovery, offer liquidity, and support market equilibrium [20]. Additionally, their trading activities have the potential to impact the actions of other market participants and market outcomes.

*Hedgers:* Hedgers engage in trading activity to manage the price risk related to their business operations [18]. To protect against potential losses from unfavorable price changes, hedging entails holding positions in the futures or options markets. Hedging traders include many participants, including producers, physical traders, refiners, and oil companies [21].

*Arbitrageurs:* Arbitrageurs use price differences between various marketplaces or contracts to generate risk-free profits [18]. They profit from temporary price differences by concurrently purchasing and selling identical items or contracts [18]. Due to factors like supply-demand imbalances, transportation costs, regional price differences, or market inefficiencies, arbitrage possibilities might occur. Significant arbitrage advantages include market efficiency, price convergence, and eliminating price differences.

### 2.1.5 Spot and Derivatives Markets

While in the spot market, the product is physically delivered immediately [22], in the derivatives market, various products guarantee delivery later [23].

*Spot Market:* Actual oil barrels are immediately bought and sold on the spot market for delivery and settlement. The delivery process includes storage, transportation, and quality control, and the time and place of delivery are usually determined in the trade agreement. While supply and demand are the fundamental drivers of the spot price, other factors also significantly affect it [24].

*Derivatives Market:* The derivatives market is a market for trading and hedging various instruments according to the price of crude oil. Market participants can hedge and gain profit from price changes without actually holding the actual commodity by using derivatives market instruments. A consensus exists for using commodity derivatives to hedge against fluctuations in commodity prices [3]. Different instruments that are offered in the derivatives market are as follows:

**Futures Contracts:** Two parties agree to buy or sell a commodity at a pre-set price, and payment occurs at a future date [23]. Futures contracts must specify their type and details like quality, size, delivery location, and time. Trades for these contracts take place on exchanges such as NYMEX, ICE, and CME. The phrase "futures price" indicates the price at which the instrument will be bought or sold when the contract expires.

**Options Contracts***:* Call and put options are the two types of options. Call options provide the buyer the right to purchase a predetermined amount of a commodity at an agreed-upon price on or before a specified future date. Put option contracts, on the other hand, provide the buyer the right to sell [25]. The absence of an obligation is the key difference between options and futures. In general, traders favor put options when they anticipate a decrease in the price of oil and call options when they anticipate an increase.

**Swaps Contracts:** Crude oil swaps are agreements between two parties to exchange cash flows based on the difference between the price of crude oil and a reference price [25]. Swaps offer protection against changes in the crude oil price without requiring actual delivery.

### 2.1.6 Major Traded Crude Oils

Over 160 different types of crude oils are traded internationally, each with a unique quality and price [4]. The most important and most widely used of these crude oils are described below.

*West Texas Intermediate (WTI):* WTI crude oil is a premium, light, and sweet crude oil refined in West Texas in the United States. WTI has an API gravity of 9.6°, which makes it light, and 0.24% sulfur content makes it sweet [26]. Because it is produced in the United States, is of outstanding quality, and is settled in US dollars, the WTI price was considered the benchmark for crude oil for years [27]. However, since 2011 WTI no longer reflects worldwide pricing, and Brent has been utilized as a benchmark [28]. Due to its high quality, its price is usually more expensive than other crude oils [26]. Its futures contracts traded on the NYMEX.

*Brent Blend:* Brent Blend oil is also light with a 38°-39° API gravity, sweet crude oil refined in Northwest Europe [26]. Compared to WTI, its sulfur concentration is greater [26]. It is called Brent Blend because it blends crude oils from various fields in the North Sea. The Brent price is currently considered the international standard for oil pricing together with WTI [29]. Its futures contracts are traded on ICE Futures Europe.

*OPEC Basket:* The Organization of the Petroleum Exporting Countries (OPEC) Basket is a weighted average of eleven crude oil produced by OPEC's member countries [22]. The basis for the weighting is the production volume and exports to the major markets [22]. It is heavier and sourer because it has a greater sulfur content and API gravity than WTI and Brent [26]. Refining crude oil becomes more difficult and expensive as the amount of sulfur in the stream increases [30]. The OPEC basket price is crucial since it serves as the benchmark for OPEC's crude oil exports and is less expensive than the prices of WTI and Brent due to its low quality.

### 2.1.7 Factors Influencing the Market

In the last few years, there's been a surge in studies focusing on the structural causes of oil price volatility [31]. Numerous factors, such as supply and demand dynamics, geopolitical tensions, economic expansions and recessions, and political issues, significantly impact the crude oil market [32]. Moreover, the price of crude oil varies from country to country due to the wide range of quality differences and the cost of transportation [33]. In the following parts, the details of the influencing factors and their impact on the crude oil market are presented.

#### 2.1.7.1 Economic Factors

Economic factors are one of the most significant influential factors of the crude oil price. Economic activity is directly linked to global oil consumption [31]. The industrial sector, which includes both manufacturing and construction, requires using energy resources to operate and expand, resulting in an increase in oil consumption during periods of economic expansion [34]. Kilian has created an index for measuring this link. The Kilian Index measures the fluctuations in industrial commodity demand associated with the cycle of the global economy [35]. In addition, many studies have tried to measure the influence of economic factors on crude oil prices by using the Economic Policy Uncertainty (EPU) index [36]–[39]. According to [40], EPU outperforms global demand, supply, and speculation in predicting crude oil price volatility.

#### 2.1.7.2 Energy Market-Related Factors

The crude oil price depends on the quantity of supply and demand, just like any product [41]. In addition to the economic factors mentioned above, crude oil demand is also affected by the price and demand for its substitutes. Natural gas and renewable energy sources such as biofuels, geothermal, and solar energy can substitute crude oil in various sectors [42].

While crude oil demand is determined mainly by economic variables and substitutes, as previously stated, numerous market participants play a vital role in crude oil supply decisions. Crude oil is split into supply from the Organization of Petroleum Exporting Countries (OPEC) and supply from non-OPEC countries [35]. Since OPEC member countries control 81% of the world's crude oil reserves [35], the decisions taken by OPEC have a significant impact on the price of crude oil. For instance, OPEC's indecisive plans to cut production cause the oil price to rise rapidly and the risk premium to increase significantly [43]. Therefore, numerous studies investigate the impact of OPEC decisions on crude oil prices [44]–[46].

#### 2.1.7.3 Financial Markets

As crude oil becomes more financialized, movements in the financial and commodities markets impact the crude oil price [47]. This effect can be investigated through four factors.

*Interest rates:* Research has indicated a negative correlation between interest rates and the crude oil price [48]. Yan et al. reported that interest rates impact oil prices as they affect the value of the US dollar [49].

*Exchange rates:* Since crude oil is traded globally through US dollars, changes in the US dollar value considerably impact crude oil prices [43]. As the US dollar depreciates, oil prices decrease, increasing oil demand [43]. For instance, in 2007, the dollar-to-euro exchange rate fell by over 10% while the global oil price increased by over 60% [50]. Numerous research studies in the literature investigate this relationship [51], [52].

*Stock market:* In the literature, it is generally accepted that stock market fluctuations and oil prices are correlated, since changes in global aggregate demand directly impact corporate profits and oil demand [35].

*Commodity market:* The crude oil price correlates to the other commodities' prices [35]. Some research has suggested that the crude oil price is related to the prices of silver [53], gold [54], and coal [55].

#### 2.1.7.4 Geopolitical Factors

Geopolitical events also influence crude oil prices [35]. Examining global oil price movements over the previous century reveals that price movements have always been closely related to the political situation in oil-producing regions [50]. Geopolitical issues can cause shortages, reduce production, and damage oil-producing areas, causing worries about the safety of oil supplies [56]. Many studies in the literature analyze the impact of geopolitical events on oil prices. For instance, a recent study suggests that the ongoing US-Iran dispute has disrupted the flow of oil worldwide and is contributing to the volatility of oil prices [57]. Another study found that geopolitical instability in the Middle East threatens oil supply, causing short and medium-term variations in oil prices [58].

The Geopolitical Risk Index is used to measure geopolitical risk in most studies in the literature [37], [59], [60]. The Geopolitical Risk Index developed by Caldara and Iacoviello is calculated based on the number of appearances of various keywords related to geopolitical risks in the leading newspapers [61]. The keywords used are grouped under various categories: tension, nuclear tension, war, and terrorism [61].

### 2.1.8 Fundamental Analysis

Analyzing and predicting price variations is the most critical requirement for profitable trading. In the literature, there are two main methods for this analysis and estimation: fundamental analysis and technical analysis [62]. Fundamental analysis investigates all elements influencing a commodity's supply and demand, whereas technical analysis investigates previous price movements to forecasts current and future prices [62]. The rest of this section details fundamental analysis and how it is used in the crude oil market.

Fundamental analysis considers all of the elements that can affect the price of an asset. These elements include the company's financial statements, managerial factors, macroeconomic factors, and investor sentiment in the case of stocks [63]. Fundamental analysis for crude oil encompasses all factors influencing its price, as detailed in Section 2.1.7. Fundamental Analysis is accomplished in three steps [6]:

- The first step is to consider all the elements that may influence the price and combine them.
- The second is to determine the sentiment of investors in the market.
- Finally, future price estimates are generated using the data from previous steps.

The development of data mining techniques and the widespread adoption of machine learning and deep learning algorithms have led to improved performance in price forecasting compared to traditional methods [63]. The following sections detail the traditional and intelligent methods used for crude oil price prediction using supply, demand, and economics-related factors and the methods used to determine investor sentiment.

#### 2.1.8.1 Prediction Models

It is essential to accurately forecast the crude oil price to determine trading strategies [64]. However, predicting the crude oil price has been challenging in forecasting research since many elements influence oil prices [65]. The crude oil price is significantly impacted by many factors, such as the balance of supply and demand and economic and political decisions and events, as discussed in Section 2.1.7. Various techniques have been employed in the literature to reveal the non-linear and non-stationary nature of crude oil price fluctuations and improve price predictions' accuracy [2]. The main types of these models are econometric models, machine learning models, and hybrid models.

*Econometric Models:* The most frequently employed econometric models in the field of crude oil price forecasting are as follows: Autoregressive Moving Average (ARMA) [66], Autoregressive Integrated Moving Average (ARIMA) [67]–[69], Generalized Autoregressive Conditional Variance (GARCH) [70]–[72], Vector Autoregression (VAR) [73], [74].

Empirical mode decomposition is a method frequently used with econometric models. This approach aims to break down complicated oil price series (non-linear and non-stationary) into many stationary components (IMFs) based on the data's intrinsic properties [75]. In studies [67] and [75], for example, time series data is decomposed into IMFs using CEEMDAN, an empirical mode decomposition approach, and the ARIMA model is used to forecast WTI and Brent crude oil prices. In the study [69], the data are decomposed into IMFs using ensemble empirical mode decomposition (EEMD), and the ARIMA model is used to forecast prices for Brent and WTI.

Since ARIMA is a linear approach, it fails to capture non-linear and non-stationary dynamics of crude oil price data [76]. Therefore, to capture the non-linear nature of the oil price, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model and its variations are widely used in literature. In [77], the authors used GARCH-N, GARCH-t, and GARCH-G models to forecast crude oil spot prices and concluded that the GARCH-N model works best for Brent and GARCH-G models for WTI crude oil spot prices. Another study uses the GARCH-M model and empirical mode decomposition (EMD) to forecast WTI prices [72]. In the study [70], the accuracy of ARIMA and GARCH models in predicting the WTI price is compared, and the GARCH model is found to perform better in daily price forecasting.

*Machine Learning Models:* As the amount of data and the complexity of variables that can influence predictions increases, the accuracy of traditional models decreases, and machine learning forecasting outperforms them, especially when dealing with non-linear problems and short-term predictions [41]. Therefore, numerous research studies use machine learning and deep learning methods to predict crude oil prices. The study [78] employed combination of XGBoost, Random Forest, and Light GBM models to forecast the crude oil price. The proposed hybrid model was found to outperform all models individually. In [78], based on VECM and the Stochastic Time-Efficient Pattern Modeling and Recognition System (STEPMRS), the authors provide a support vector machine-based (SVM) ensemble model for forecasting monthly WTI spot prices. Their proposed model outperforms individual models. In another study, a modified linear regression-based machine learning model was used to predict Brent prices [79]. In Study [80], the success of six different machine learning models, such as Multiple Linear Regression, K-Nearest Neighbor Regression, Random Forest, XGB, Light Gradient Boosting Machine, and CatBoost, in predicting WTI spot price is compared, with the conclusion that LGBM provides the best results.

Neural networks and machine learning models are commonly utilized in time series forecasting [81]. Several studies use various types of neural networks to predict the crude oil price. For instance, Long Short-Term Memory (LSTM) was used in the study [81], Artificial Neural Networks with Whale Optimization Algorithm (WOA) in the study [82], Wavelet Neural Network (WNN) in the study [83], Deep (or hierarchical) Multiple Kernel Learning (DMKL) in study [84], Functional Link Artificial Neural Network (FLANN) in study [85], Multiple Wavelet Recurrent Neural Networks (MWRNNs) and BPNN in study [86] and MLTSM with Mahalanobis and Z-score transformations in study [1].

*Hybrid Models:* In addition to research that employs machine learning and econometric/statistical models to forecast crude oil prices, some studies use hybrid methods that combine the benefits of these approaches. Using news feeds and historical price data, the study [2] aims to estimate the daily WTI futures price. First, the authors extracted the sentiment scores of the news using TF-IDF (Term Frequency - Inverse Document Frequency), then decomposed the time series data into IMFs using Ensemble Empirical Mode Decomposition (EEMD). Finally, the Seagull Optimization Algorithm (SOA) is utilized to optimize the prediction model parameters, namely Gated Recurrent Unit (GRU). In another study, time series data were decomposed using Multivariate EMD (MEMD), and Brent spot price was forecast using Linear Regression, Back Propagation Neural Network (BPNN), Extreme Learning Machine (ELM), and Random Factor Functional Link Network (RVFL) [87]. In another hybrid study, Variational Mode Decomposition (VMD) was used for decomposition, and Bidirectional Gated Recurrent Unit (BiGRU) was used for prediction. [88] In [65], the authors proposed AGESL, a hybrid framework that employs a fully connected neural network to combine ARIMA's average price forecast, GARCH's volatility forecast, and LSTM's price forecast.

#### 2.1.9 Sentiment Analysis

Sentiment Analysis analyzes people's attitudes, beliefs, and feelings toward others, events, or concepts [89]. With the widespread use of the Internet, people use social media to share their thoughts about other people, products, and events, leading to the growing popularity of sentiment analysis [90]. Sentiment analysis is given a text as input, while the output is either 'positive,' 'negative,' or 'neutral' with a numeric value in its fundamental form. In this view, sentiment analysis is a classification problem with each category representing a sentiment [91].

Document, sentence, and aspect are the three levels of sentiment analysis [92]. The document level evaluates a document's positive or negative content as a whole, the sentence level evaluates the sentiment of individual sentences, and the aspect level attempts to achieve the idea without structural limitations [92].

Sentiment analysis has three main methods: machine learning approach, rule-based approach, and lexical-based approach [92]. Rule-based Sentiment Analysis (RBSA) uses pre-defined rules or patterns based on prior language knowledge, intuition, or expertise in a particular field to determine the sentiment of a text [92]. While lexicon-based approaches use the pre-polarity lexicon to figure out the sentiment analysis, machine-learning techniques usually use a classifier to categorize texts based on their sentiments [93].

In recent years, with the development of deep learning models and the emergence of Large Language Models (LLMs), the tasks that can be done with sentiment analysis have become more diverse. For instance, ABSA (Aspect-Based Sentiment Analysis) is a subset of sentiment analysis that concentrates on detecting sentiment toward particular elements or characteristics within a text [94]. On the other hand, Multifaceted Analysis of Subjective Text (MAST) is another subset of sentiment analysis that covers a variety of specialized assignments that focus on distinct phenomena of sentiment or opinion [94]. With this diversification, it is now possible to execute various tasks, such as detecting hate speech against diverse groups, investor sentiments towards financial products, and product sentiment based on consumer reviews.

Financial sentiment analysis differs from general sentiment analysis in that it typically attempts to forecast how markets will react to the information offered in the text [95]. Financial sentiment analysis dates back to the 1980s. The Bag of Word (BOW) technique was utilized in the first examples of sentiment analysis in the finance field [96]. This lexicon-based method derives a sentiment score from the total of positively and negatively labeled words [96]. Machine learning-based approaches have been utilized in sentiment analysis as machine and deep learning technology has advanced. The transformer model is a tool that has revolutionized language modeling tasks by implementing self-attention mechanisms that capture contextual dependencies [96]. This model was first introduced in 2017 in Vaswani et al.'s research titled "Attention Is All You Need" [97]. Following the appearance of the Transformer model, language models that may be used for sentiment analysis began to develop and became capable of performing various tasks. In 2018, Google released BERT (Bidirectional Encoder Representations from Transformers) [98], and OpenAI released GPT (Generative Pre-Trained Transformer) [99]. Following the emergence of these large language models, models such as GPT-2 [100], GPT-3 [101], and GPT-4 [102] that are trained with more data and perform better, as well as domain-specific LLMs such as FinBert [95], BloombergGPT [103], BioBert [104], SciBert [105], and LawGPT [106] that are fine-tuned with domain-specific data sets, have emerged.

To evaluate the impact of investor sentiment on crude oil prices, sentiment analysis can be applied to crude oil trading. Studies show a strong relationship between crude oil prices and investor sentiment [107]–[109]. Crude oil price forecasting studies try to measure investor sentiment with data collected from news sources [2], [88], [110], [111], social media [112], and Google Trends data [87], [113], [114] and try to increase forecast accuracy by incorporating the data obtained into forecasting models. Many studies used FinBert [96], a variant of BERT [98] that is trained using financial data to handle natural language problems in the finance domain [88], [115]. It should be noted, however, that FinBert fails to comprehend the dynamics of the oil market and differentiate the sentiment of a particular event from its effect on the price of crude oil [116]. For example, whereas FinBert classifies a disruption in the crude oil supply chain as an adverse event, a decline in crude oil supply actually raises the price [116]. Kaplan et al. developed CrudeBert, a FinBert variant trained on crude oil-related data, to address this issue [96], [116].

### 2.1.10 Technical Analysis

Although there is no evidence, it is believed that technical analysis has been employed since the beginning of the trade [117]. The basis of technical analysis was analyzing the relationships between recorded or remembered prices in these trades and attempting to predict future prices [117]. Charles Dow established the groundwork for modern technical analysis in the 1800s [2]. Even though he has no books, he has published several articles in the Wall Street Journal detailing his stock market calculations [7]. After Dow's death, S.A. Nelson combined his publications and created the "Dow Theory" in his book The ABC of Stock Speculation [118]. The usage of technical analysis for trading and the emergence and spread of new technical indicators with the help of sophisticated calculations that become possible with advancing technology has continued since Dow's Theory. Technical analysis is a means of understanding the current state of the markets through charts and indicators and predicting future price patterns and trends in this manner [7]. Chart analysis, pattern recognition analysis, seasonality and cycle analysis, and automated technical trading systems are techniques used in technical analysis [119].

Academics have been skeptical about technical analysis, which is growing in popularity and use among practitioners [117]. Two notable studies contradicting technical analyses are the Random Walk Theory and the Efficient Market Hypothesis. Louis Bachelier established the foundations of the Random Walk Theory by arguing that the basis of the random walk was irrationality [120]. According to this theory, the price is determined randomly, and historical data provides no insight into the future. On the other hand, according to Eugene Fama's Efficient Market Hypothesis, the current price of an asset shows whole market information [121]. There are three forms of The Efficient Market Hypothesis. In the weak form, it is considered that the current asset price reflects all historical data [122]. As a result, using historical data to profit from the current market is pointless. The semi-strong form states that asset prices show all publicly available information and that the market immediately adjusts to new information [122]. This indicates that fundamental analysis cannot create an advantage in the market. According to the strong form, the asset price reflects all information known or unknown to the public [122]. On the contrary, more modern theories, such as Noisy Rational Expectation Theory and behavioral theories, imply that technical analysis can be beneficial due to market noise and market players' irrational behaviors [119]. Defenders and skeptics of these theories have conducted extensive research on the profitability of technical analysis in various markets. In their study published in 2004, Park and Irwin investigated 92 papers that attempted to measure the profitability of technical analysis in different markets [119]. While it was concluded that technical analysis provided positive returns in 58 of these studies, it was found that it was not profitable in 24, and mixed results were obtained in 10 [119]. In another review conducted in 2017, studies using technical analysis in the stock market were examined. Among the 85 studies reviewed, it was found that 79 of them supported the technical analysis, while 6 of them did not [123].

Park and Irwin's another study investigated the profitability of technical trading rules for 17 futures markets, including crude oil futures, from 1985 to 2004 [124]. The profitability of almost nine thousand rules belonging to 14 technical trading systems was tested. It was found that these rules are not profitable when transaction costs, and data snooping effects are considered [124].

Another study on how technical trading rules performed in the crude oil market assessed the performance of the 7846 technical trading rules proposed by Sullivan et al. in the WTI futures market and United States Oil (USO) fund [125], [126]. Although the results showed that profitable trades could be made even when transaction costs are included, the persistence analysis concluded that small profits might be gained in some periods, and there is no lasting quality in implementing the rules [125].

Another study examined whether technical trading rules are profitable and whether different crude oil markets confirm the Efficient Market Hypothesis using daily spot price data from the WTI, Brent, and XLE markets between 1999 and 2021 [127]. The duration of the Covid-19 pandemic was investigated independently to quantify the influence of the pandemic on crude oil markets. More than 14 thousand technical trading rules have been tested on this data, and White's Reality tests have been applied to measure the data snooping effect. While the technical trading rules did not generate an additional profit for the WTI market between 1999-2021 and the Brent market between 1999-2019, however, did result in profitable outcomes in the Brent market in the Covid-19 period [127].

Another study seeks to determine if futures prices can be used to forecast spot market prices in four commodity markets, including crude oil, and whether technical trading rules are profitable in these spot markets [128]. The results show that futures returns are a good predictor of spot returns, and moving average and range break trading rules provide statistically significant positive profit for the crude oil market [128].

Unlike previous studies, Ji's study attempted to determine if the technical trading rules for the Chinese crude oil market are profitable by utilizing intraday data rather than daily data [129]. When transaction costs, and data snooping effects are included, profitable results may be achieved in in-sample periods; however, persistence analysis revealed that no profitable set of rules could be employed [129].

With the advancement of technology, the indicators employed in technical analysis have gotten more complex and are now being incorporated into intelligent systems. For this reason, it is essential to look at studies that use technical analysis along with intelligent systems when questioning the profitability of technical analysis.

According to studies that employ technical indicators in econometric models, technical indicators provide statistically and economically significant results in predicting the price of oil. Study [130] attempts to predict the WTI Crude Oil closing price using OLS Regression with the help of 18 technical indicators based on the moving-average rule, momentum rule, on-balance volume rule, and various macroeconomic data. The Principal Component analysis revealed that while 18 technical indicators significantly predict crude oil futures returns, macroeconomic variables cannot [130]. The authors consider investor sentiment to be one of the factors in the technical indicators' ability to predict the price of oil [130]. By incorporating real-time data, another study improved the study of Ying and Yang [130], [131]. The authors concentrated on density forecast rather than point forecast since they evaluated the predictability of oil prices from investors' perspectives. They attempted to predict the returns of WTI futures using constant coefficient and time-varying parameter regressions [131]. They employed ten technical indicators based on moving averages and nine macroeconomic factors as explanatory variables [131]. The findings led to the conclusion that technical indicators produce forecasts that are statistically and economically significant [131]. In another study, 18 technical indicators recommended by Ying and Yang [130] and 18 macroeconomic indicators those that are often employed in the literature to predict stock return, those that reflect the overall state of the economy, and those that depict oil supply and demand balances were used as explanatory variables to predict WTI spot prices [132]. The authors used an iterative econometric forecasting method, unlike other studies, and according to the results obtained, technical indicators provide economically and statistically significant predictive power on crude oil prices, and they work better than macroeconomic indicators [132]. Another study examines the ElasticNet and Lasso models' ability to predict the oil price utilizing a broad set of predictors that includes 14 macroeconomic factors and 18 technical indicators proposed by Ying and Yang [130], [133]. Consistent with the findings of the studies above, it was concluded that technical indicators have more predictive power than macroeconomic data.

When we examine the studies that incorporate technical indicators into machine learning models, in the study [88], macroeconomic indicators, sentiment analysis from news, and six technical indicators, the simple moving average (SMA), exponential moving average (EMA), moving average convergence-divergence (MACD), relative strength index (RSI), Williams %R, and stochastic oscillator (%K) were used to forecast the daily WTI price using the FinBERT-VMD-Att-BiGRU hybrid model, and it was concluded that technical indicators increase the forecast accuracy. Another study evaluated the accuracy of predicting the WTI price of 8 different models utilizing 18 technical indicators suggested by Ying and Yang [130], and it was concluded that there was a relationship, although not a strong one, between crude oil price and technical indicators. [37] Another study analyzed the success of the technical indicators proposed by Ying and Yang [130] in estimating the crude oil price, and it was concluded that the technical indicators outperform the macroeconomic indicators in terms of their predictive power. [133]

## 2.2 Decision Support Systems

A Decision Support System (DSS) is a computer-based system supporting the users' decision-making by providing them with necessary data, knowledge, models, and analytical tools [134]. DSSs are generally classified into five categories: data-driven, communication-driven, document-driven, model-driven, and knowledge-driven decision support systems [135]. As technology advances, Intelligent Decision Support Systems based on artificial intelligence techniques have emerged [134]. These systems substantially assist in tackling challenging problems that are typically real-time and include vast volumes of distributed data by utilizing AI technologies such as Artificial Neural Networks and Intelligent Agents [134].

### 2.2.1 Financial Decision Support Systems

Decision Support Systems began to be used in financial decision-making in the 1970s, emphasizing financial planning, portfolio management, and banking [136]. Due to technological advancements that have transformed the financial sector, modern financial decision support has embraced many innovative technologies [136]. Modern technologies such as reinforcement learning [137], artificial intelligence [138], [139], and data mining [140] are commonly applied in financial DSSs in the literature. When reviewing the DSSs in the trading studies, most systems were proposed for stock trading. One of the earliest studies to present a decision support system for stock market investors was published in 1997 by Liu and Lee [141]. Their study solely used technical analysis indicators such as Stochastics, Relative Strength Index, Moving Averages, and Trend Lines. The DSS, called StockAdvisor, comprises two subsystems that manipulate charts and data and make intelligent conclusions based on experts' domain knowledge using rule-based reasoning. The system's prototype was constructed and tested using Hong Kong stock market data. Another study stated that stock market forecasts based only on technical indicators are inadequate and that qualitative data such as political events or investor sentiment should also be included in the models [142]. Their proposed decision support system consists of three main components: the first component brings together all the factors that affect the stock market, the second component includes fuzzy logic that aims to incorporate investors' domain expertise into the system, and the last component combines the qualitative data obtained in the second part with technical indicators using a neural network. However, in this study, the influencing events for the stock market are decided by a group of experts and incorporated into the system at regular intervals, which may result in inaccurate judgements [142]. In another study, a DSS for stock trading was presented that employs an Open-High-Low-Close chart and a Japanese Candlestick chart to determine the stock's trend, as well as the Williams %R indicator to determine whether the stock is overvalued or undervalued [140]. The DSS also generates trading signals based on various trading rules and back tests them using historical data. Furthermore, an AI system for predicting stock prices based on diverse financial inputs has been proposed. The study's limitation is that it does not consider the investor sentiment or market related news [140].

When looking at the decision support systems offered for stock trading, it is clear that there are studies that use only technical analysis [143], [141] fundamental analysis [138], both [140], [144], and sentiment analysis [145], but there is no comprehensive study that integrates all of these concepts.

### 2.2.2 Crude Oil Trading Decision Support Systems

There are few DSS studies for crude oil trading in the literature. One of the pioneering studies conducted in 1999 proposed a decision support system based on the Blackboard Model, which consists of crude oil evaluation component based on expert opinion, inventory evaluation, and transportation cost evaluation components, which include the presentation of various data, and economic evaluation component, in which price and demand are estimated using multiple econometric methods [146]. It is proposed that all data in the decision support system be collected in a single segregated database and that experts determine which data groups are used in which sub-components and econometric forecasts [146]. Another study conducted in 2022 uses historical WTI and Brent daily prices to forecast the daily return with the rolling window method, then develops a trading signal based on its prediction and presents the risks and returns of the proposed trading strategy to the user using historical data [147]. This study also allows users to specify the parameters of the proposed prediction model themselves. However, in both of these studies, the decision of crude oil trading was not considered holistically by employing different analyses, they only use different methods to forecast crude oil prices.

As a result of the extensive review of the literature, knowledge gaps were identified as follows:

1. No study in the literature proposed a comprehensive decision support system for crude oil trading.
2. No decision support system combines all the approaches that need to be considered in crude oil trading, such as fundamental, technical, and sentiment analysis.

# CHAPTER 3

# RESEARCH METHODOLOGY

## 3.1 Research Aim & Objectives

The primary objective of this research is to propose a decision support system for individual traders seeking to profit in crude oil markets. The main focus is on individual investors with limited market knowledge and difficulty accessing the necessary data. In this context, the research questions were determined as follows:

- What factors affect the crude oil price?
- How can a framework for a system to support trading decisions for crude oil be proposed and implemented?

Based on the research questions, the research objectives of the study are as follows:

- To systematically specify the data requirements of the crude oil decision support system.
- To propose a framework to support decisions on crude oil trading by utilizing technical, fundamental, and sentiment analysis.
- To develop the implementation of the proposed framework to support crude oil trading decisions.

## 3.2 Research Methodology

Design Science Research (DSR) is a paradigm that creates innovative artifacts consistent with the prior literature [148]. It mainly applies to information systems (IS) research, where creating effective artifacts is crucial [149]. Therefore, DSR is the research methodology used in this study. DSR can be implemented in various methods; however, this study uses the methodology proposed by Peffers et al. [149]. The Design Science Research Methodology (DSRM) includes a mental model for presenting and evaluating DSR in IS, a process model for performing DSR, and assures that the research process is consistent with previous literature [149]. The steps of the suggested framework for implementing DSR are shown in Figure 1. A research entry point, problem identification and motivation, the determination of objectives, design and development, demonstration, evaluation, and communication are the seven steps that make up this framework [149].

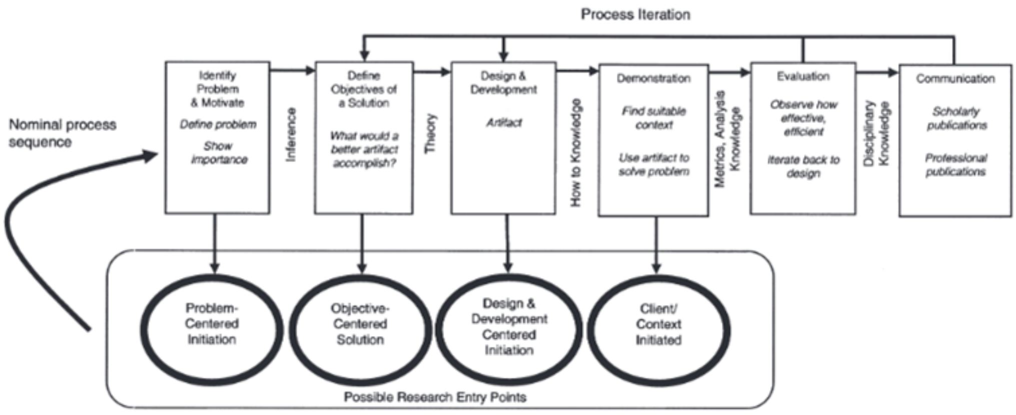


Figure 1: DSRM Process Model [149]

### 3.2.1 Research Entry Point

The research entry point should be determined according to the specific study context and objectives [149]. The four research entry points are problem-centered initiation, goal-centered solution, design and development-centered initiation, and customer/context-centered initiation, as seen in Figure 1 [149]. A problem-centered initiation is used in this study. This entry point can be used when researchers see a problem or shortcoming in a particular field and want to create a solution [150]. Since a comprehensive literature review revealed the absence of any studies on the topic, this research was conducted to solve this problem.

### 3.2.2 Identify Problem and Motivate

Problem identification and motivation is the initial phase in DSRM, where researchers identify a problem or challenge that needs to be solved [151]. This phase is essential as it lays the groundwork for the entire research process. The difficulties individual traders in the crude oil market encounter are the primary focus of this study, which also brings together several indicators and data that can be used to make profitable trades. Its primary goal is to propose and implement a decision-support system that individual crude oil market traders might utilize to understand market dynamics better and perform profitable trades.

### 3.2.3 Define Objectives of a Solution

Defining the objectives of the solution to the problem indicated in the first stage is the second step of the DSRM [149]. The objectives can be quantitative or qualitative [149]. In this study, qualitative objectives were identified. These objectives and the advantages of the proposed decision support system are described in detail in the following sections.

### 3.2.4 Design and Development

The authors define this stage as artifact creation [149]. Models, constructs, methods, or instantiations to be created in line with the defined objectives should be the output of this phase [149]. The proposed decision support system is an artifact of this study. This artifact is a model designed and developed with the objectives mentioned above.

### 3.2.5 Demonstration

The demonstration, the fourth step of the DSR Methodology, can be accomplished using experimentation, simulation, case studies, proofs, prototypes, or other suitable methods [149]. A prototype of the proposed trading decision support system has been developed in this study. This prototype demonstrates the applicability of the proposed model.

### 3.2.6 Evaluation

Evaluation is a crucial part of the DSR Methodology [149]. It entails assessing the results of DSR to ensure they are efficient and beneficial [152]. There are different evaluation methods. The methods proposed by Hevner et al. are shown in Table 1 [153]. Experimental evaluation methods were utilized in this study.

Table 1: Evaluation Methods [153]

| Category | Evaluation Method |
|---|---|
| **Observational** | Case Study<br>Field Study |
| **Analytical** | Static Analysis<br>Architecture Analysis<br>Optimization<br>Dynamic Analysis |
| **Experimental** | Controlled Experiment<br>Simulation |
| **Testing** | Functional (Black Box) Testing<br>Structural (White Box) Testing |
| **Descriptive** | Informed Argument<br>Scenarios |

### 3.2.7 Communication of Research

Another crucial phase in DSRM is the communication of research, which entails explaining the research's methodology, findings, and implications to various stakeholders. This can be achieved through academic papers, a thesis, and participation in conferences and symposiums. The details of the research for DSR are given as proposed by Hevner et al. in this thesis which is the primary communication of the research method [153]. One of the methodologies proposed in this artifact was also presented at the following conference:

- IISEC (https://iisec.tbdakademi.org.tr) - III. International Informatics and Software Engineering Conference, “A Comprehensive Big Data Framework for Energy Markets: Oil and Gas Demand Forecasting” by Alp Bayar, Burcu Koç, Mert Onuralp Gökalp, Baran Özden, Yiğit Yeldan, P. Erhan Eren, Altan Koçyiğit

# CHAPTER 4

# PROPOSED DECISION SUPPORT SYSTEM FRAMEWORK

## 4.1 Introduction

This chapter proposes a decision support system for the crude oil trading problem. The proposed decision support system corresponds to the artifact creation stage in the Design Science Research Methodology [153]. The proposed DSS aims to answer the two research questions of this study:

1. What factors affect the price of crude oil?
2. How can a framework for a system to support trading decisions for crude oil be proposed and implemented?

A systematic literature review was conducted following the guidelines proposed by Kitchenam [154] to answer the first research question and identify the factors affecting the price of crude oil to be used in the proposed DSS, and 558 different data sets influencing the price of crude oil were compiled and presented. The remainder of this chapter contains detailed information regarding the proposed decision support system and its components for answering the first stage of the second research question.

As described in Chapter 2, a decision support system should integrate data, modeling, intelligence, and visualization to enhance decision-making. Also, in the same chapter, it is highlighted that predicting the price of crude oil is critical in making trading decisions and that various factors influence this price. In addition to making the price forecast by considering all influencing factors (fundamental analysis), it is crucial to consider the past volatility of the price to understand future patterns and trends (technical analysis). Furthermore, investor sentiments significantly impact the market and should be considered when trading (sentiment analysis). Figure 2 presents the proposed decision support system in light of this information.

The Real-Time Data Collection and Integration Layer is the first component of the proposed DSS and is responsible for the real-time collection, preprocessing, and integration of all datasets from various sources into a unified database for the crude oil trading decision-making process. In the Modelling Layer, the Fundamental Analysis Component contains predictive AI models for crude oil price-related forecasts, while the Sentiment Analysis Component contains various NLP algorithms. The Intelligence and Visualizations Layer includes the user interface, modeling predictions, and technical analysis recommendations.

The datasets that can be used in the proposed DSS and the steps of the Systematic Literature Review used to gather these data are explained in the remainder of this chapter. Then, additional data that can be used in the DSS is discussed, and an architecture for collecting and integrating all of this data from various sources is presented. Following that, the features of the Modelling Layer and the Intelligence and Visualizations Layer are discussed.

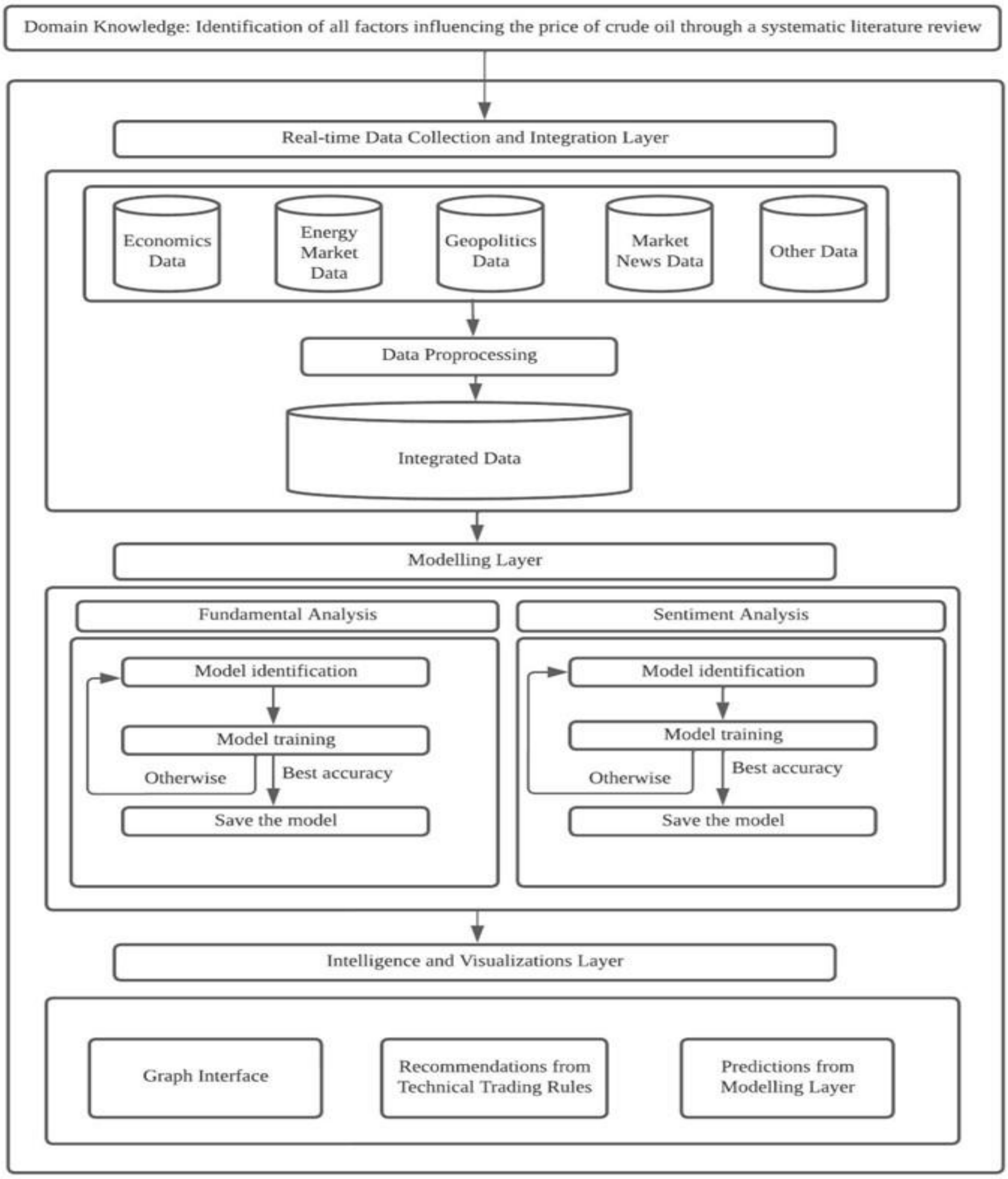


Figure 2: Proposed crude oil trading decision support system

## 4.2 Real-Time Data Collection and Integration Layer

In the proposed decision support system, combining the influencing factors to forecast the crude oil price is crucial. Influencing factors of the crude oil price, such as macroeconomic data, news flows, supply and demand statistics, and so on, are published at various intervals, such as real-time, daily, weekly, and monthly by multiple sources. For this reason, it is essential to collect and format these data consistently to gain insight from the data and modeling. Big data technologies are suggested to gather, preprocess, and store data in real-time and historically in the proposed system to address this issue. In the following section, the steps of the systematic literature review used to determine all the factors impacting the crude oil price for predictive modeling are discussed, and the influencing factors are presented. Following that, the details of the data required for the other components of the decision support system are provided. Finally, the details of the suggested big data infrastructure to collect and preprocess this data are described.

### 4.2.1 Data for Predictive Modelling

Following Kitchenham's suggested procedures [154], a systematic literature review is conducted to determine the influencing factors of crude oil prices. For this reason, the research question of the review was determined as follows:

- Which data sources are employed in studies that use AI techniques to forecast the price of crude oil?

*Search Strategy:* The search terms were decided to be oil type (“Crude”, “WTI”, “Brent”), price type (“price”, “spot”, “futures”), and task (“forecast”, “estimate”, “predict”) to make the search as comprehensive as possible.

*Search String:* The search string for the initial literature search is determined as (Crude OR oil OR WTI OR Brent) AND (price OR spot OR futures) AND (forecast OR estimate OR predict OR predictability OR predictive OR estimating OR forecasting OR predicting OR prediction OR estimation).

*Search Scope:* The preliminary search used the Scopus and Web of Science databases. The review was conducted in February 2023, and studies published after 2010 were included in the analysis.

*Initial Search and Selection Process:* The predetermined search string was used to search the publications' abstracts, titles, and keywords in the relevant databases, and 30,844 studies were accessed, 19,789 from Scopus and 11,055 from Web of Science. Applying the elimination criteria shown in Table 2, 7138 studies in Scopus and 6822 studies in Web of Science were reached; 9,548 studies were identified as the initial pool when duplicates were removed from these studies.

Table 2: Selection Process

| Selection Process | Databases |
|---|---|
| Language: English | Scopus |
| Date ≥ 2010 | Web of Science |
| Publication: Journal | |
| Version: Final | |
| Relevance: Abstract, Title, Keyword | |

*Primary Studies:* The second elimination was accomplished by reading the titles and abstracts of 9548 studies determined in the initial pool. As a result of this process, it was decided that 235 studies were suitable for the scope of this review. After the full-text reviews of the accepted articles, 47 studies were determined as the primary studies. As a result, full-text reading of the 47 studies shaped the results of this review.

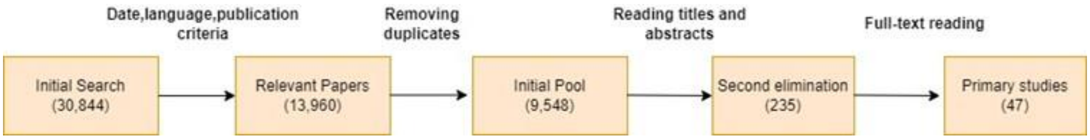


Figure 3:Systematic literature review

### 4.2.2 Systematic Literature Review Findings

This study examined crude oil price forecasting studies using machine learning and hybrid methods. 235 studies identified in the second elimination are grouped according to their data sources and methods. Details of the grouping can be seen in Figure 4. While 215 of these studies are doing price forecasting, 12 are forecasting price trends, six are forecasting return, 1 is crush, and 1 is term structure.

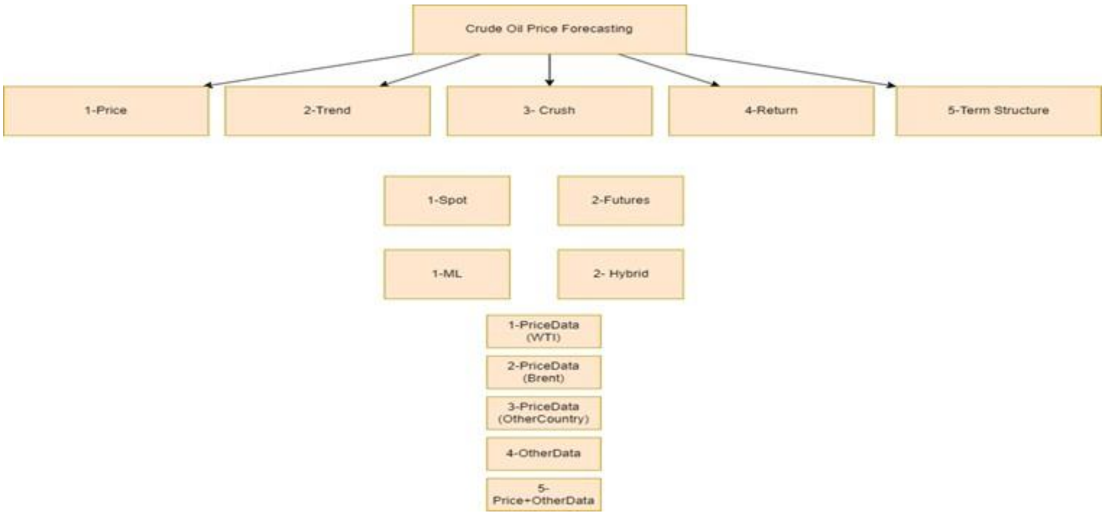


Figure 4: Grouping of the studies in second elimination

Since the primary purpose of this review is to determine the different data sources used in forecasting major crude oil prices in the literature, 47 studies that make forecasting using various data sources have been selected as the primary studies. Only historical oil price data was used to predict in 168 of 215 price forecasting studies. It is insufficient to estimate solely on price data because many factors affect the price of crude oil, including supply-demand-related factors, the status of the financial markets, and the overall economy. Table 3 has detailed information on the primary studies. 39 of the 47 primary studies employed spot price forecasting, seven employed futures price forecasting, and one employed both spot and futures price forecasting. While daily price forecasts are made in 17 of these studies, there are also forecasts made for other time intervals, such as weekly in 5 studies, monthly in 21 of them, annual in 3, and monthly to varying periods in 1 study.

Identifying and collecting data sets from various sources that affect crude oil prices is one of the most fundamental elements of the proposed decision support system. The data sources we gathered during our systematic literature review on studies that predict crude oil prices using multiple data sources will be employed in the proposed crude oil decision support system.

The data utilized in the price predictions for the various periods in the literature are released at varying intervals. For example, daily forecasting studies often rely on weekly and daily data. Since the vast bulk of macroeconomic data is released monthly, it is also possible to use a wider variety of data for monthly price forecasting. Depending on the duration of the estimation, it is possible to choose which data sources we have identified from the literature may be utilized in the modeling.

The complete list of 558 various data sets utilized to forecast the WTI, Brent, Forties, Dubai, Oman, and USA F.O.B. cost of OPEC crudes on a daily, weekly, monthly, and annual basis is provided in Appendix A. Each data set is further separated into 12 categories: consumption, orders and inventory, economic uncertainty, energy outlook, financial and commodity markets, geopolitical factors, housing, industrial production, interest and exchange rates, labor market, macroeconomic indicators, money and credit, and prices. These categories can be grouped under three primary headings. These include geopolitical, energy market, and economics data. Further in this chapter, details of the data sources and relevant categories will be presented.

Table 3: Primary Studies

| **Study** | **Oil Market** | **Time Scale** | **Models** |
|---|---|---|---|
| [1] | WTI spot | Daily | Mahalanobis and Z-score transformations with MLSTM |
| [2] | WTI futures | Daily | EEMD-SOA-GRU-MLR |
| [155] | WTI spot | Monthly | Stacked denoising autoencoders (SDAE)-NN |
| [86] | WTI/Brent spot | Yearly | Multiple wavelet recurrent neural networks (MWRNNs)-BPNN |
| [87] | Brent spot | Weekly | SED-MEMD-LR/BPNN/ELM-RVFL |
| [156] | WTI spot | Daily | RF-XGBoost-Light GBM |
| [88] | WTI futures | Daily | FinBERT-VMD-Att-BiGRU |
| [78] | WTI spot | Monthly | SVM-VECM and STEPMRS |
| [80] | WTI spot | Daily | MLR, KNN, RF, XGB, LGBM, CATB, DNN |
| [37] | WTI spot | Yearly | GCT-SFA-LR, ANN, SVR, and RF |
| [41] | WTI spot | Monthly | BMA-LSTM, Slab-LASSO-LSTM and GLMNET-LSTM |
| [85] | WTI spot | Daily | Functional Link Artificial Neural Network (FLANN) |
| [157] | WTI spot | Daily | Ensemble –ANFIS-PSO |
| [158] | WTI spot | Monthly | Comparison of VAR, VEC, and ANN models |
| [159] | Simple average of WTI, Brent, and Dubai spot | Monthly | PSO-Radial basis function neural network |
| [160] | WTI futures | Daily | TextBlob-LDA-SeaNMF-RF,SVR,<br>Arima, Arimax |

Table 3 continued:

| | | | |
|---|---|---|---|
| [110] | WTI spot | Daily | CEEMDAN-LSTM_att-ADD with the news sentiment index (TFIDF) |
| [38] | WTI futures | Monthly | Sentiment analysis-VAR |
| [113] | WTI spot | Weekly | BPNN, SVM, RNN, LSTM, GRU |
| [84] | WTI, Brent, Forties, Dubai, and Oman spot | Daily | Deep (or hierarchical) multiple kernel learning (DMKL) |
| [161] | WTI spot | Monthly | Wavelet Neural Network (WNN) |
| [65] | WTI spot | Daily | AGESL (ARIMA, GARCH, ODEE, LSTM) |
| [162] | WTI spot | Monthly | K-means-KPCA-KELM |
| [163] | WTI spot | Weekly | LR, SVR, BPNN, ELM, RVFL |
| [133] | WTI spot | Monthly | ElasticNet and LASSO |
| [164] | Simple average of WTI, Brent, and Dubai spot | Yearly | VAR and ANN |
| [82] | WTI spot | Monthly | ANN-WOA |
| [39] | WTI/Brent spot | Monthly | Lasso-TVP Regression |
| [165] | Brent spot | Daily | Ridgeregression, LASSO, SVR, BPNN, and RF |
| [59] | WTI, Brent, Dubai spot | Monthly | Bayesian Symbolic Regression |
| [166] | WTI, RAC spot | Monthly | ElasticNet, LASSO, Ridge with Huber Loss Function |
| [111] | WTI spot | Monthly | BPNN, SVM, and MLR |
| [167] | WTI spot | Monthly | Google index-driven decomposition ensemble (GIDDE)-VAR |
| [168] | WTI spot | Monthly | GA-SSA-ANFIS |
| [35] | WTI spot | Daily | Lasso |
| [60] | Brent spot | Monthly to ranging times | 23 different ML and econometric models |

Table 3 continued:

| | | | |
|---|---|---|---|
| [169] | WTI/Brent spot | Monthly | VMD-GASVM and VMD-GABP |
| [114] | WTI spot | Daily | VMD-SVM, RF, LSTM, BPNN, LSTM&BP |
| [170] | Brent spot | Monthly | NN-Genetic Algorithm |
| [79] | Brent spot | Weekly | Modified linear regression-based ML algorithm |
| [171] | USA F.O.B. cost of OPEC spot | Monthly | Adaptive neuro-fuzzy inference system (ANFIS) |
| [64] | WTI futures | Daily | Comparison of various econometric and ANN models |
| [172] | Brent spot | Monthly | Intelligent model search engine |
| [49] | WTI spot/futures | Monthly | LLE-RNN/LSTM |
| [173] | WTI spot | Daily | CNN-LDA-RF, SVR, LR |
| [81] | WTI futures | Weekly | LSTM |
| [115] | WTI futures | Daily | LDA-TextBlob-FinBERT-CNN,LightGBM |

#### 4.2.2.1 Economics Data

The oil price and the overall economy have a relationship in two directions. Numerous research studies have examined this relationship [174]–[176]. As a result, it is crucial to include economic factors while making oil price predictions. Various economic indicators are categorized and described in the following paragraphs.

*Consumption, Orders, and Inventory Data:* This category includes information on consumer spending patterns, which is crucial for determining economic performance. The Consumer Sentiment Index is among the most significant of these data. Through a variety of consumer surveys and an analysis of how consumers will shape their purchasing decisions in response, this index aims to gauge the degree of optimism consumers have regarding the overall state of the economy. Other information on consumer spending patterns and amounts, such as order, stock levels, and personal expenditures, are also available in this category. According to [60], global real economic activity is one of the critical predictors of crude oil price; this data can provide crucial information on the price of crude oil because they track the state of the economic conditions.

*Economic Uncertainty Data:* The policy-related economic uncertainty index, which aims to gauge how unpredictable people believe the market will be following changes in economic policies, appears in this category. For key economies, including those of Canada, China, Russia, the UK, and the US, the policy-related economic uncertainty index may contain important information about the state of the global economy and, consequently, the trajectory of the price of crude oil. According to [37], West Texas Intermediate (WTI) oil price volatility can be predicted with highly predictive accuracy using EPU indices. This association exists because crude oil suppliers frequently hoard inventories to reduce risks associated with global EPU, which causes greater volatility in crude oil prices [37].

*Financial and Commodity Markets Data:* This category includes a variety of prices, ratios, and indicators that reflect the state of the significant stock and commodity markets. The price of crude oil and the prices of other commodities are related. As a result, the data set contains the prices of commodities, including gold, silver, iron ore, copper, coal, and natural gas. In particular, since crude oil and gold are both denominated in dollars, the prices of these two commodities have a strong correlation [1]. For example, oil price increases will result in inflation, raising gold prices. Additionally, the price of crude oil may also be impacted by alternative commodities that can be used as a substitute for it, such as natural gas, coal, and renewable energy sources [155].

Before 2007, there was no discernible relationship between oil prices and the financial market or exchange rates. The financialization phenomenon did, however, significantly correlate the price of crude oil with the price of financial assets and exchange rates following this year [80]. As oil products become more financially integrated, the relationship between the oil and financial markets has grown stronger [37]. Stock markets are significantly affected by short-term and long-term crude oil shocks, illustrating the positive correlation between stock prices and oil prices [156]. The data set contains information on the functioning of the financial markets of other significant economies in addition to the American financial market. For instance, China, one of the world's largest oil consumers and economies, impacts oil price variations [115].

*Housing:* Since the housing industry and the economy's overall health are closely related, changes in one can significantly impact the other. There are numerous sources of this relationship. For instance, there is an uptick in residential sales when individuals can save, and the economy is performing well. In addition, through mortgage financing, the financial sector and the housing market are closely linked. These factors lead to the inclusion of housing-related data in the presented data sets, such as newly built houses and average house prices.

*Industrial Production:* The volume of industrial production offers crucial information regarding the demand for oil because crude oil is the primary energy source and essential raw ingredient for many sectors. For instance, in the petrochemical industry, crude oil is transformed into different chemicals, polymers, and synthetic materials used to make various goods, such as construction materials, medications, textiles, and fertilizers. Additionally, transporting raw materials, intermediate commodities, and finished goods over great distances is frequently necessary for industrial manufacturing. Several forms of transportation, including ships, trucks, trains, and airplanes, are propelled by oil-based fuels. For these reasons, industrial production is a factor that influences and is used to predict the price of crude oil, and there are studies in the literature that try to increase the accuracy of oil price estimation by combining industrial production and capacity utilization data that reflect the speed of industrial development [37].

*Interest and Exchange Rates:* The foreign exchange market significantly influences oil prices because the US dollar is widely used to gauge the global economy's health [1]. Because worldwide crude oil is priced and traded in US dollars, a weak dollar tends to drive up oil prices [37]. The dollar exchange rate is one of the most significant predictors of the oil price and is included as the predictor in the majority of the studies reviewed. In addition to the US Dollar Index, the exchange rate between the dollar and other major currencies, such as the Pound and Euro, has been included in the studies.

The demand for crude oil rises during economic expansion; hence there is a significant correlation between the state of the world economy and the price of crude oil. For this reason, various interest rates that provide information on the state of the economy should be considered when studying the variables affecting the price of crude oil. One of these, and one of the most significant interest rates in the American economy, the Federal Fund Rate, impacts the core elements of the overall economy, including employment, growth, and inflation [157]. Investor sentiment is another factor in the correlation between crude oil prices and interest rates and influences oil prices. Changes in interest rates can significantly affect decisions regarding investments and borrowing costs since they impact risk appetite and investor confidence.

*Labor Markets:* The state of the labor market might reveal the economy's overall health. A low unemployment rate generally indicates a healthy labor market with stronger consumer purchasing power. It is anticipated that oil demand will rise during these times, and hence indicators relating to the labor market can provide information regarding the oil price.

*Macroeconomic Indicators:* As previously stated, the state of the world economy is one of the key factors influencing the demand for crude oil. Therefore changes in oil prices can be explained using a variety of macroeconomic factors [37]. According to numerous research, using macroeconomic data improves the accuracy of crude oil price forecasting models [37]. These factors include the GDP of major economies, the PMI Index, the Chicago Fed National Activity Index, and other data that provide information about the economy's direction. For example, global economic activity and growth, a key driver of oil demand, are reflected in world GDP growth [164].

*Money and Credit:* Most studies in the literature conclude that the money supply and economic growth are positively correlated [177]. The quantity of money available for investors to spend and invest rises with an increase in the money supply, and economic growth accelerates. For this reason, data on the money supply, such as the M1 and M2 money stocks, reserves of the depository institutions, and the monetary base, is included in studies in the literature. Total consumer loans, motor vehicle loans, and real estate loans are also included in the data set since they are tools for monetary policy and indicators of the state of the economy.

*Prices:* There is a bidirectional relationship between inflation and crude oil prices. During times of increasing inflation expectations, investors in the crude oil market may demand higher compensation for holding oil contracts which could lead to an increase in crude oil prices. In addition, rising crude oil prices could cause inflation by raising the cost of production. For this reason, inflation-related data such as Consumer Price Index and Producer Price Index are included in the data set.

#### 4.2.2.2 Energy Market Data

The primary elements influencing the long-term direction of oil prices have always been variations in supply and demand [41]. In general, demand from industrialized countries and supply from exporting countries determine the price of crude oil [80]. According to several studies, oil futures prices are driven by four factors: oil inventory, oil aggregate supply, oil aggregate demand, and oil speculative demand [115]. The demand for other petroleum products should be considered to calculate the demand for crude oil. As a result, the data set also contains information on the demand for other petroleum products, such as tight oil, gasoline, and fuel oil. The demand for fossil fuels and renewable energy sources that can replace crude oil is another factor influencing its demand.

#### 4.2.2.3 Geopolitical Data

Political circumstances also influence the price of oil [1]. According to [60], various economic and geopolitical developments over the past 50 years have shaped the real price of crude oil. For instance, crude oil prices have significantly changed due to political events like the Iran-Iraq War and the 9-11 attack. Geopolitical Risk Indexes for various nations are therefore included in the data set. In the literature, multiple studies demonstrated that GPR influences investor attitudes and trading decisions, which results in oil price changes [37].

All data sets acquired from the literature that affects the crude oil price can be included in the forecasting models together or separately. The forecast range is critical when selecting data for modeling. Furthermore, it is important to give traders flexibility to choose from various data sources for input to forecasting models.

### 4.2.3 Data for Sentiment Analysis

Many studies in the literature have concluded that there is a significant relationship between investor sentiment and market prices [88]. Investor sentiment is measured using a variety of data sources. In crude oil price forecasting studies, news sources are the key data sets for sentiment analysis [2], [38], [88], [110], [160]. Aside from that, some studies use data from Google Trends [80], [87], [162], [163] and numerous social media sources [112]. Data for studies that employ news sources for sentiment analysis was gathered from sources such as oilprice.com [2], [111], [115], [155], [165], investing.com [160], [173], the Financial Times [38], Reuters [165], Google News [88], and The New York Times [65]. One or more of these data sources could be used in the proposed decision support system. For instance, news and social media information can be used together to measure investor sentiment. Sentiment analysis findings can become more realistic when data is collected from multiple sources.

### 4.2.4 Other Data

In addition to the data sets stated above, recommended for crude oil price forecasting and investor sentiment modeling, different data sets should be included in the proposed decision support system for data analysis and technical analysis. To calculate the technical indicators and trading rules used in technical analysis, it is necessary to use the asset's daily opening, closing, lowest and highest price, and volume data [178]. Calculating these indicators and trading rules for various time frames is also critical for optimizing trading decisions by assessing short, medium, and long-term trends [179]. As a result, it is suggested in the proposed decision support system to use real-time data to execute necessary calculations for various time frames ranging from 1 minute to 1 month based on the needs of traders.

### 4.2.5 Data Integration

Sections 4.2.1, 4.2.2, 4.2.3 and 4.2.4 provide extensive detail regarding the data that could be used in the proposed decision support system. These data are provided from various sources, in multiple forms, and over various periods. For instance, the data to be used for fundamental analysis is published in different periods, such as real-time like gold price and US dollar index, daily like Morgan Stanley Capital International (MSCI) World Index, and Baltic Exchange Dirty Tanker Index, weekly like US crude oil stock, or monthly like U.S. refinery utilization rate. Furthermore, news feeds, crude oil spot, and futures pricing data, which may be used for technical and sentiment analysis, should be received in real-time. One of the most critical tasks in ensuring that decision support systems provide satisfactory outcomes is to efficiently collect this heterogeneous data set and transform it into the same format used in modeling and analysis [180]. To overcome this problem, the proposed decision support system needs to be integrated into a big data infrastructure. The following sections discuss big data, how it can be integrated into the proposed crude oil decision support system, and what challenges it can solve.

#### 4.2.6 Big Data

As a result of the increased usage of the Internet and smart devices, as well as the advancement of technology, millions of individuals are producing massive volumes of data [181]. This data is called Big Data and cannot be processed using conventional methods and tools due to its volume, variety, velocity, value, and veracity [182]. Volume is the most essential and distinguishing feature of big data, as well as the primary reason why traditional methods cannot be used to process big data [183]. Volume can be defined as the amount of data [182]. Variety refers to diverse data collected from various sources, whereas velocity means the high frequency of data collection and processing [184]. The term value relates to the benefits of big data, whereas veracity refers to the trustworthiness of the collected data [182].

Because traditional approaches are insufficient for processing big data due to the above characteristics, new methods, and techniques are required to improve decision-making using big data [182]. Distributed, parallel, in-memory, and cloud computing technologies enable Big Data [185]. Due to the high volume feature of big data, it is impossible to perform data storage and processing operations on a single machine. Distributed computing solves this problem by processing data across multiple nodes or machines, increasing scalability and efficiency [186]. On the other hand, parallel computing can be analogized to a divide-and-conquer approach. It refers to dividing a task into smaller incremental steps that can be completed simultaneously, thereby increasing computational speed and productivity [186]. Although distributed computing was a primary focus of most early big data research, cloud computing and big data are now frequently used in tandem [185]. This is because cloud computing offers essential shared computing resources for overcoming the difficulties associated with the storage, processing, and distribution of Big Data and improving the benefits of distributed and parallel processing [185].

New data management tools have been developed to address the challenges associated with big data. The most important is Apache Hadoop, an open-source distributed system infrastructure and software platform [184]. It was built by Doug Cutting and Mike Cafarella in 2004 [187]. Hadoop includes the Hadoop Distributed File System (HDFS) and MapReduce. HDFS is a dependable storage component of the Hadoop architecture that partitions and stores data on multiple machines to provide fault tolerance [188]. MapReduce is a programming model of the Hadoop architecture for processing big data [188]. The Hadoop ecosystem includes different software tools for various purposes, such as storing, processing, and visualizing big data such as Hive, Pig, Kafka, and Spark.

### 4.2.7 Using Big Data in Crude Oil Trading DSS

The proposed decision support system requires collecting, integrating, and processing various data types from multiple data sources at different time scales. These processes are critical to the efficient functioning of the decision-making process. Infrastructure is vital to addressing these challenges, so integrating big data infrastructure is recommended.

In the proposed big data infrastructure's first component, the necessary data must be collected and ingested from external data sources. Apache Kafka, an open-source distributed event streaming platform, can be used to perform these tasks [189]. Kafka is commonly used to develop real-time data applications that receive streaming data from numerous sources [190]. Kafka is based on producer-to-consumer communication, with data collected and transferred in categories called topics [189]. Kafka is an optimal message center for real-time analysis of big data streams due to its high throughput, reduced latency, robustness, and fault tolerance [191].

The structured, semi-structured, or unstructured data collected through Kafka should be converted into the same format before the modeling process. This requires a system that can preprocess both real-time and batch data. Apache Storm, an open-source real-time computation system, can be employed for preprocessing [189]. Because trading decisions can sometimes be made in seconds, real-time data must reach the system with low latency. Regarding performance, it is reasonable to say that Apache Storm outperforms similar technologies in terms of low latency [192]. Storm can also perform batch processing, although its main strength is real-time analytics. Because the proposed decision support system incorporates historical data, batch processing with Apache Spark is also an option. Since it has a library for building machine learning models with big data called MLlib, [193], it can be utilized in modeling with historical data.

The proposed decision support system also requires a database that supports a wide range of data structures to accommodate the high volumes of data used and the results generated. Therefore, a distributed NoSQL database such as Apache Cassandra can be used. Since it provides high availability, eliminating the possibility of system failure and ensuring eventual consistency, it is a suitable database for the proposed decision support system [194].

The literature was utilized to identify data sources that can be used for the real-time data collection and integration layer, which is the initial component of the proposed decision support system. These data sources come in a variety of formats and are released by a variety of sources. The importance of employing big data architecture for the efficient functioning of this process and integrating real-time and historical data from external sources with the DSS has been discussed above. Using the recommended big data architecture, mechanisms that are arranged based on data release times and continually update the data in the system can be built.

## 4.3 Modeling Layer

Variations in the price of crude oil are significant for traders and policymakers due to the impact of oil price changes on global economic conditions [155]. As a result, several academic studies are attempting to forecast crude oil prices [195]. However, as described in Section 4.2.1, crude oil prices are affected by various factors, making crude oil price forecasting challenging. To reflect the non-linear and non-stationary nature of crude oil prices, researchers have used a wide range of forecasting models in addition to enriching data sources [1]. The models utilized might be classified as econometric, machine learning, and hybrid. This study suggests to usage of nonlinear machine learning methods because typical linear methods fail to capture the nonlinear dynamics of oil price time series [1].

Due to the wide range of variables that influence the price of crude oil, financial market developments, including statements of country leaders, central bank chairman, oil companies, and other major oil market participants, as well as social media posts by investors, can also contain important information about the direction of the crude oil market [195]. As a result, sentiment analysis has become a widely used technique in crude oil price forecasting literature [195].

One of the most significant features of the modeling layer is that it is not dependent on any particular model. Different models may produce superior results when data types and periods are changed. As a result, it is proposed to retrain the models used in the proposed system for different types of data provided to the system and use the model that produces more accurate results. This section details the fundamental and sentiment analysis components of the modeling layer of the proposed decision support system, the framework of the modeling to be performed in these components, and the possible intelligent models that can be employed.

### 4.3.1 Fundamental Analysis Component

The main objective of the fundamental analysis component in the modeling layer of the proposed crude oil decision support system is to execute models that forecast the crude oil price using appropriate data from the real-time data collection and integration layer. A wide variety of models are used for crude oil price forecasting in the literature, and these models can be categorized into three main groups: econometric, machine learning, and hybrid [156]. Since traditional econometric techniques for conventional time series forecasting do not provide satisfactory outcomes for oil price forecasts [88], [78], this study suggests using non-linear machine learning or hybrid models to capture the non-linear and non-stationary nature of the crude oil price series. Data mining techniques and artificial intelligence have enabled the development of new methods for analyzing and predicting crude oil prices [78]. Table 3 contains detailed information on the models employed in studies utilizing machine learning and hybrid models for crude oil price forecasts. Detailed information about the models widely used in the literature and can be used in the proposed decision support system is given in the following sections.

#### 4.3.1.1 Machine Learning Models

Machine learning models in time series forecasting have become increasingly popular due to their ability to simulate complex features such as non-linearity and volatility [81]. Support Vector Machine, Extreme Gradient Boosting (XGBoost), and Light GBM (LGBM) are machine learning models typically employed in crude oil price forecasting literature [80].

*Support Vector Machines (SVM):* Support Vector Machine (SVM) is a machine learning model built on Statistical Learning Theory [78]. In its original state, SVM seeks an optimal hyperplane for separating data into multiple classes [78]. Although it originated to solve classification problems, it is also widely used for regression problems [113]. The formulization of the SVM is as follows:

$$min_{w,b,\xi} \frac{1}{2}||w||^2 + C\sum_{i=1}^{N} \xi_i$$

$$s.t.\, y_i \left( \left( w.\phi(x_i) \right) + b \right) \geq 1 - \xi_i \,, i = 1, \ldots.., N$$

$$\xi \geq 0, \;\; i = 1, \ldots\ldots, N$$

where x = $\phi(x_i)$ represents the transform form the original space $x_i \epsilon \boldsymbol{R}^n$ to Hilbert space, w represents the slope of the hyperplane in Hilbert space, C represents the coefficient of the penalty term and $\sum_{i=1}^{N} \xi_i$ represents penalty term that means cost of training [78]. SVM is one of the most commonly employed models in crude oil price forecasting studies [2], [155], either individually [113], [173], or in combination with other models [78].

*Extreme Gradient Boosting (XGBoost):* XGBoost is a machine learning model proposed in 2016 and used for classification and regression problems [80]. XGBoost is a model built on the Gradient Boosting Decision Tree (GBDT) approach [156]. This method involves the initial training of a tree, followed by the subsequent training of a second tree based on the results of the trained model with the actual value, and the process is repeated until it reaches a predetermined value or constraint [156]. The outcome of XGB is the summation of the prediction values from all Classification and Regression Trees (CARTs) and is formulated as follows:

$$\gamma = \sum_{t=1}^{N} f_m(X)$$

where N represents number of trees, $f_m$ denotes each CART tree, and $\gamma$ is the prediction result [80]. XGBoost is also commonly employed in crude oil price forecasting research, individually [80] or as part of hybrid models [156], and has been shown to outperform traditional models.

*Light Gradient Boosting Machine (LGBM):* Light GBM, like XGBoost, is a Gradient Boosting Decision Tree (GBDT) model [156]. It can provide the accuracy of XGBoost but is faster due to its structure that develops leaf-wise trees [115]. LGBM's estimation function is shown below:

$$y_m = \sum_{t=1}^{N} f_m\,(X)$$

where $f_m(X)$ represents regression tree, and N is the number of regression trees. It has succeeded in many forecasting challenges, including crude oil price forecasting [196]. It has been utilized in numerous studies on crude oil price forecasts in the literature [80], [115], [156].

#### 4.3.1.2 Neural Networks

Neural networks are algorithms based on the human brain and built to identify patterns. They consist of input and output layers and hidden layers [197]. These levels are made up of nodes. Figure 5 depicts the structure of neural networks in their simple form. The nodes' primary role is to design the learning process by assigning significance or insignificance to the information to be learned based on the model's input. They accomplish this by assigning different weights to the inputs. The sum of these weights is then processed through an activation function to determine how much influence this signal has on the outcome. Several activation functions include Sigmoid, Hyperbolic Tangent, Rectified Linear Unit (ReLU), and Softmax [198]. As each activation function has distinct features, the appropriate one can be determined based on the characteristics of the problem. Learning takes place due to the weights allocated to neurons and the activation function, which determines the effect of these weights on the final output.

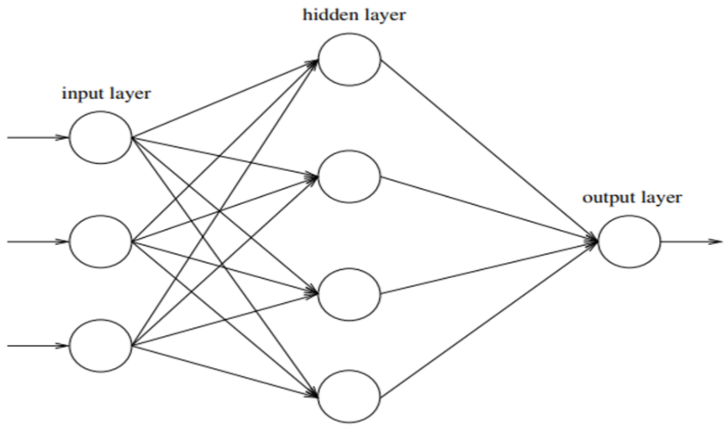


Figure 5: Simple neural network [197]

Neural Networks are non-linear models frequently used in many applications such as classification, regression, and natural language processing [199]. Using neural networks to enhance crude oil price forecasting has become increasingly popular among researchers in recent years [1]. A wide variety of artificial neural network models such as Long Short-Term Memory (LSTM) [1], [41], [65], [110], Stacked Denoising Autoencoder Neural Networks [155], Multiple Wavelet Recurrent Neural Networks [86], Bidirectional Gated Recurrent Unit (BiGRU) [88], Functional Link Artificial Neural Network [85], Radial Basis Function Neural Network [159], Wavelet Neural Network [83], Backpropagation Neural Network (BPNN) [111], [114] are used in the literature for crude oil price prediction.

#### 4.3.1.3 Neural Networks for Time Series Forecasting

Recurrent Neural Networks (RNN) are the most commonly utilized neural network type for time series forecasting [200]. The reason for this is the capacity of RNNs to hold previous information and incorporate it into the prediction model [200]. Although there are other varieties of RNNs, including Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Bidirectional RNN, [201] Long Short-Term Memory (LSTM) is one of the most widely used ones in crude oil price forecasting studies [1], [41], [49], [65], [110], [114].

LSTM was proposed in 1997 with a research paper called "Long Short-Term Memory" [202]. The fundamental purpose of LSTM is to tackle the problem of vanishing gradients in RNNs. This can be described as the fact that as the number of layers in a neural network rise, the value of the derivative product approaches a value converge to zero, and so its effect is almost negligible [202]. The LSTM comprises two subsystems: cell state and hidden state, and the input, output, and forget gates. The input gate controls the cell state and the forget gate, while the output gate manages the hidden state [201]. The formulation of LSTM is as follows:

$$i_t = sigmoid(W_i x_t + U_i h_{t-1} + b_i)$$

$$f_t = sigmoid(W_f x_t + U_f h_{t-1} + b_f)$$

$$o_t = sigmoid(W_o x_t + U_o h_{t-1} + b_o)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_c x_t + U_c h_{t-1} + b_c)$$

$$h_t = o_t \odot \tanh(c_t)$$

The LSTM architecture ensures that information is retained over a long period by providing a mechanism for determining what information needs to be remembered and for how long [201]. The effectiveness of LSTM in time series forecasting has led to its widespread application in a wide range of fields [81].

### 4.3.2 Sentiment Analysis Component

The Sentiment Analysis Component's primary function is to scan news feeds and analyze the impact of identified news on crude oil prices. A wide range of models have been utilized to observe this effect. Methodologically, as detailed in Section 2.1.9, sentiment analysis can be classified into three categories: rule-based, lexicon-based, and machine learning-based [92]. When we look at which models have been used for sentiment analysis using news text in crude oil-related studies, it is seen that models from various categories, such as Term Frequency-Inverse Document Frequency (TF-IDF) [2], [110], [113], FinBert [88],[115] dictionary-based approach (computing positive and negative words) [38], event extraction with Open Domain Event Extraction (ODEE) algorithm [65], Valence Aware Dictionary and Sentiment Reasoner (VADER) [165] and CNN [111], [173] were employed. In most of these studies, sentiment analysis results are given as input to the models constructed for crude oil price forecasting [2], [87], [88].

Due to the outstanding outcomes acquired thanks to the Transformer Model detailed in Section 2.1.9, Large Language Model-based sentiment analysis is becoming increasingly popular nowadays [94]. Based on transformer architecture, the current generation of sentiment classifiers can achieve high accuracies of 97.5% or higher [203]. Moreover, since LLMs are both trained with massive datasets and have fine-tuned versions with domain-specific data makes them appealing for sentiment analysis [94]. Despite these advantages, the success of LLMs in financial sentiment remains controversial [204]. FinBert is the most widely used LLM for financial sentiment analysis [88]. When performing sentiment analysis, FinBert [95], a variation of the Bert model trained with financial data, returns a value between [-1,1]. Texts can be classified as positive, negative, or neutral in this manner. Although trained with financial data, FinBert does not fully capture the crude oil market dynamics [96]. CrudeBert, a FinBert variant trained on crude oil-related news, attempts to address these problems [96]. The impact of 46,000 crude oil-related news headlines on the price of crude oil according to the law of supply and demand was labeled, and 60 percent of this data was used in FinBert's fine-tuning. According to the test results, CrudeBert predicts the impact of crude oil market news on price more accurately than FinBert [96]. Therefore, in the proposed system, the CrudeBert model can be used to measure the impact of news flows on crude oil price. In addition, the GPT versions, which are recently becoming very widely used, can also be used to observe this effect. No study in the literature uses GPT versions for crude oil-related sentiment analysis. However, a very recent study using news feeds and ChatGPT-4 to measure stock market returns concluded that GPT-4 achieved outstanding results [205].

### 4.4 Intelligence and Visualization Layer

The last layer of the proposed decision support system, the Intelligence and Visualization layer, uses data and models from the other layers to provide data analysis, recommendations, and data visualization to support the crude oil trading decision. One of the critical characteristics of the layer is the ability to make real-time recommendations in support of the trading decision, as the use of real-time data is recommended, and the models operating with this data will be continuously updated. Section 2.1.10 discusses how historical price values are crucial in trading decisions and forecasts based on the factors influencing crude oil prices. In this regard, this layer should be able to calculate and visualize technical indicators and make recommendations based on technical trading rules. Furthermore, data visualizations are crucial components of a decision support system that aid decision-making. Data visualization facilitates the identification of anomalies in data, tracking temporal changes, and identifying clusters and data shifts [206]. In the remainder of this section, Intelligence and Visualizations layer is explained in detail under two separate headings.

### 4.4.1 Intelligence Layer

The most crucial component of the Intelligence layer is the calculation of technical trading rules and the derivation of recommendations based on these rules. As explained in Section 2.1.10, technical trading rules are a method of technical analysis. Qualitative and quantitative are two distinct types of technical analysis. Qualitative technical analysis involves examining charts and attempting to identify trends in the data, while quantitative technical analysis involves analyzing time series data to create trading signals [207]. Due to the numerous techniques and indicator combinations developed by traders and analysts over the years, technical trading rules are almost infinite. Traders usually employ five types of technical trading rules [208]. These are oscillator trading rules, filter trading rules, moving average trading rules, support-resistance trading rules, and channel breakout trading rules [208]. The parameter selection for these rules is crucial as different parameters may produce different returns [209].

*Moving Average Trading Rules:* The technical analysis literature refers to moving average rules as one of the most widely used and commonly discussed trading rules [126]. Moving averages are often obtained from the average closing price over a certain period and are calculated using a variety of window sizes (e.g., 5, 10, or 100 days) [210]. In its basic form, the moving average trading rule states that when the current price is greater (or equal to or less than) the average price over the preceding trading period, it constitutes a buy signal (or sell signal or neutral signal) [211].

*Oscillator Trading Rules:* These rules change depending on the indicator utilized but examine whether an asset is overbought or oversold [208]. The Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), Stochastic Oscillator, and Commodity Channel Index (CCI), Rate of Change (ROC) are among the popular oscillator technical indicators.

*Filter Trading Rules:* They try to follow trends by issuing a buy or sell signal when the price rises or falls by a certain percentage [209].

*Support-Resistance Trading Rules:* In these trading rules, support and resistance points are determined around the price based on various criteria, and it is concluded that the price will continue its movement in the same direction when it surpasses these points [208].

*Channel Breakout Trading Rules:* These trading rules establish a time-variant price channel, which, when crossed, indicates movement in the same direction [209].

There are studies to determine the profitability of technical trading rules in the crude oil market. The details of these studies are covered in Section 2.1.10. Some studies suggest that technical trading rules are profitable, while others conclude that they are not profitable when considering transaction costs. However, there is no doubt that even if traders do not follow the signals of technical trading rules, they can still use them to understand the market's direction. As a result, the proposed decision support system must incorporate technical trading rules appropriate for traders' requirements and trading strategies. Moreover, technical trading rules may also produce signals that differ [140]. As a result, in the proposed system, traders must see the profitability of these rules over different periods through back-testing and make decisions accordingly.

### 4.4.2 Visualizations Layer

The modeling results, recommendations based on technical analysis, and data visualizations must be provided to the user through an interface in the decision support system proposed to support traders' trading decisions. The primary function of this layer is to transfer information acquired from other layers of the system to users in a user-friendly manner. The details of the user interface for different components of the proposed system are explained in detail in the rest of this section.

#### 4.4.2.1 Fundamental Analysis

The modeling results conducted in the fundamental analysis component of the modeling layer should be displayed to the users. These models should consider all factors influencing crude oil prices identified in a systematic literature review. Daily, weekly, monthly, and even annual price forecasts can be generated by selecting appropriate data types based on the traders' needs. In addition to point price forecasts for various time intervals, modeling results such as price trend forecasts, change of the price increase/decrease forecasts, and so on should be displayed on the user interface. Consequently, forecasts of the point price, the amount of price change, and the direction of change should be provided to users with models developed utilizing appropriate data types for different time scales.

Moreover, independent observation of the impact of influencing factors in crude oil price forecasting can provide valuable market information to traders. Therefore, one of the most important features of the proposed system is that it provides the user with the flexibility to select the data to be input to the price forecasting models and to present the results of these models. To this end, the user interface can display the datasets to be included in the models in different categories such as economic, energy market related, geopolitical, technical indicators and so on, allowing traders to compare the forecast results generated with the data they choose and learn more about the market trends and underlying factors. In addition, it is also important to provide users with the necessary visualization mechanisms to observe the relationship between the data given as input to the forecasting models and the crude oil price, which is the target variable.

#### 4.4.2.2 Sentiment Analysis

The findings derived from the models operating in the sentiment analysis component should be included in the user interface of the proposed system. Daily news feeds for crude oil and other market components should be analyzed on a news-by-news basis, and users should be provided with estimates of the impact of each news on the price of crude oil. Furthermore, the literature indicates that the correlation between oil news and investor sentiment is continuous, which implies that the combination of the news on a particular day and the prior several days influences investor sentiment daily [2]. Models that employ investor sentiment measured by the news as input to crude oil price forecasting models use different approaches to measure daily sentiment scores to observe this effect. One study, for example, used the Cumulative Sentiment Score to solve this issue [2]. This score computes the cumulative score so that the influence of more recent news grows exponentially. A daily sentiment score is derived in another study by averaging each news-based sentiment score [160]. In addition to news-based sentiment analysis in the proposed system, hourly, daily, weekly, and monthly cumulative sentiment should be calculated with different methods used in the literature, such as linear summation, weighted summation, and averaging, and presented to users. It is also crucial for the proposed system to provide users with access to news of a specific day and the modeling results of this news.

#### 4.4.2.3 Technical Analysis

In addition to the outcomes of the modeling layer, it is crucial to provide users with recommendations, graphics, and data analysis about technical analysis in the user interface. As explained in Section 2.1.10, the profitability of technical analysis is a matter of debate in the literature. For this reason, it aims to support the trader's decision-making process and illustrate historical price data by using technical analysis methods such as charts, indicators, and trading rules together rather than providing precise results. A combination of multiple approaches outperforms a single method [212]. The remainder of this section examines the modules associated with technical analysis in the user interface.

*Candlestick Charts:* Candlestick charts are a widespread technique in technical analysis. These charts, which display an asset's opening, closing, high, and low values, provide helpful information for price analysis. Figure 6 shows the structure of the candlestick chart. Candlestick pattern recognition plays an essential role in individual transactions [213]. It allows traders to estimate the current market value of assets and whether the current selling/buying momentum will continue [213].

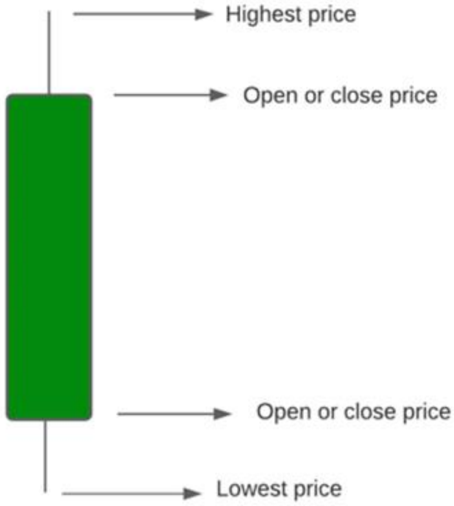


Figure 6: Candlestick chart

Many studies in literature have focused on candlesticks. Researchers have employed various techniques to determine price trends, including trend continuation and reversal indicators such as Morning Star and the Evening Star [214]. In the proposed crude oil trading decision support system, the candlestick chart of the crude oil price for different periods should be presented to the users.

Another chart that should be offered to users along with the candlestick chart is the volume chart [7]. Volume indicates the trading activity in the market at a specific time [7]. According to studies in the literature, volume can be used to gauge the market's reaction to new information and changes in investor sentiment [215].

*Technical Indicators:* Technical indicators are another technique utilized in technical analysis. The following parts provide detailed information on the most often utilized technical indicators in crude oil trading [216]. Using multiple technical indicators improves the accuracy of the signal generated [212]. As a result, all of these technical indicators, and any indicators that can be determined based on the demands of the traders, should be made available to users in the proposed system.

**Moving Average***:* Moving average is one of the most widely used indicators in technical analysis. Moving averages are classified into several categories, but the simple moving average is being used in this study. Its formulation is as follows:

$$MA_n = \frac{1}{m} \sum_{i=n-m+1}^{n} p_i$$

**Relative Strength Index:** The Relative Strength Index (RSI) is a momentum oscillator measuring the price movement rate [217]. The Relative Strength Index (RSI) is commonly used with a data period of 14 days. It provides a range of values between 0 and 100, with an overbought value above 70 and an oversold value below 30 [217]. The mathematical formulation of RSI used in this study is as follows:

$$Average\ Gain = \frac{Sum\ of\ gains\ (positive\ changes)\ over\ the\ past\ t\ days}{t}$$

$$Average\ Loss = \frac{Sum\ of\ Losses\ (negative\ changes) over\ the\ past\ t\ days}{t}$$

$$RS = \frac{Average\ Gain}{Average\ Loss}$$

$$RSI = 100 - \frac{100}{(1 + RS)}$$

**Bolinger Bands:** Bollinger bands are lines drawn at various standard deviations above and below a simple moving average of a price [218]. The volatility of the asset influences the difference between the bands. Proximity to the upper band may indicate overbought conditions, and proximity to the lower band may indicate oversold conditions.

**Moving Average Convergence Divergence (MACD):** The Moving Average Convergence Divergence (MACD) is a widely used momentum technical indicator. It measures the degree of convergence and divergence using two moving averages, 12 days EMA and 26 days EMA [219]. The mathematical formulation of the MACD used in this study is as follows:

$$MACD = \sum_{i=1}^{n} EMA_k - \sum_{i=1}^{n} EMA_d$$

$$EMA_n = EMA_{n-1} + \alpha(p_n - EMA_{n-1})$$

Where k is 12, d is 26, n is time, $p_n$ is price of the asset at that time [219].

**Stochastic Oscillator:** It is an indicator used to measure the relative position of an asset's closing price to prices over a specified period. It is represented by two lines,

%K and %D [212]. In the context of this research, the calculation of a Stochastic oscillator is performed as follows:

$$\%K = 100 * \frac{C_t - L_t(m)}{R_t(m)}$$

$$\%D = \frac{\sum_{i=t-n}^{t} \%K_i}{n}$$

where $C_t$ is closing price, $L_t(m)$ is lowest price of the last t days and $R_t(m)$ is the price range.

**Fibonacci Retracement:** Fibonacci retracement is a method that utilizes horizontal lines to identify potential levels of support and resistance on a price graph using Fibonacci sequences. In this study, Fibonacci retracement levels, which are widely used in the literature, are 23.6%, 38.2%, 61.8%, and 78.6%, are used [220].

**Williams %R:** Williams %R is an oscillating momentum indicator. This indicator, which has a value between 0 and -100, usually indicates overbought between 0 and -20 and oversold between -80 and -100 [221]. The mathematical formulation is as follows:

$$\%R_t = \frac{Highest\ high - Close_t}{Highest\ high - Lowest\ Low}$$

where highest high is the maximum price for the specified period, lowest low is the minimum price, and close is the closing price at the end of the period.

**Momentum:** The momentum indicator is a simple indicator used to assess a price change's direction and strength. The following is the momentum calculation utilized in this study:

$$Momentum_t = Close_t - Close_{t-n}$$

**Rate of Change:** ROC is a technical indicator used to evaluate the rate of change in an asset's price over a specified period. The rate of change is calculated as follows:

$$ROC_t = \frac{Close_t - Close_{t-n}}{Close_{t-n}}$$

**Average True Range:** ATR is a technical indicator used to measure the variations in the price of an asset over a specified period [222]. The ATR used in this study is calculated as follows:

$$True\ Range_t = Max[(High\ Price_t - Low\ Price_t), (High\ Price_t - Closing\ Price_{t-1}), (Low\ Price_t - Closing\ Price_{t-1})$$

$$ATR = \frac{\sum_{i=1}^{n} True\ Range_i}{n}$$

**Commodity Channel Index (CCI):** The CCI is an oscillating momentum indicator. It indicates the market is overbought when its value is above +100 or oversold when below -100 [221]. The CCI used in this study is calculated as follows:

$$P_t = \frac{Low\ price + High\ price + Closing\ price}{3}$$

$$CCI_t = \frac{1}{0.015} - \frac{P_t - SMA(P_t)}{Mean\ Absolute\ Deviation\ (P_t)}$$

**On-Balance Volume:** On-Balance Volume (OBV) is a technical indicator used to measure buying and selling pressure. On days when the price increases, it adds the volume and vice versa. The mathematical formulation is as follows:

$$C_t > C_{t-1} \Rightarrow OBV_t = OBV_{t-1} + Volume_t$$

$$C_t < C_{t-1} \Rightarrow OBV_t = OBV_{t-1} - Volume_t$$

$$C_t = C_{t-1} \Rightarrow OBV_t = OBV_{t-1}$$

*Technical Trading Rules:* Section 4.4.1 provides detailed information about technical trading rules. Buy and sell signals generated by these rules should be displayed to the users. The selection of parameters for these rules is critical since various parameters may generate different results [209]. Therefore, these rules should be calculated according to the traders' needs for different periods. Furthermore, the profitability of technical trading rules in the crude oil market is controversial in academia, with some research concluding that it is difficult to make statements of overall profitability [124] while others provide periodic profitability [125], [127], [129]. As a result, it is crucial to back test the technical trading rules presented to users with historical data to demonstrate their profitability over a period of time specified by users. In this way, users can see which technical trading rules are more profitable at different time periods when the price shows a similar pattern and make trading decisions accordingly.

# CHAPTER 5

# PROTOTYPE

This chapter provides details of the prototype of the proposed system. The prototype aims to answer the second part of the second research questions of this study:

- How can a framework for a system to support trading decisions for crude oil be proposed and implemented?

The prototype of the proposed system is explicitly applied to traders in the WTI futures market. Using the prototype, traders can see forecasts of the WTI Spot price and the WTI Front Month Futures price for the next day and one week ahead and make their trading decision to stay or exit the market accordingly. They can also use signals from hourly, four-hourly, daily, and weekly candlestick charts, technical indicator charts, and technical trading rules to support their short-term and long-term trading decisions for the WTI Front Month Futures. They will also be able to utilize sentiment analysis to foresee market direction by considering investor sentiment and how daily news flows affect crude oil prices. The proposed system and the prototype aim to provide traders with a holistic perspective by combining fundamental, technical, and sentiment analysis. In addition to the speculators who trade for profit, individuals or organizations who want to sell physical products, hedge, or determine their strategy based on the price of crude oil can also gain valuable insights from the proposed system. Moreover, the proposed system, in which the factors affecting the price of crude oil are identified through an extensive systematic literature review, can also be used as part of a more comprehensive algorithmic trading system.

As stated in Section 4.1, the proposed decision support system is an artifact of the DSR process built by merging diverse concepts, tools, and technologies from literature to solve a real-world problem. Implementing the proposed system to demonstrate its effectiveness in solving this problem is also necessary [153]. Therefore, an implementation of the proposed system for the WTI Futures market has been presented. As detailed in Section 2.1.6, WTI is one of the most traded crude oils in the world, and its price has been used as a benchmark to determine the price of crude oil for years due to its high quality. WTI futures are traded on the NYMEX. While market participants trade WTI physically in the spot market, they frequently trade in the futures market to hedge and profit. To profit from the market, traders must monitor the direction and variations in crude oil prices, considering the elements that can influence the price. The projection of an upcoming increase/decrease in the price is the among the most important indicators for traders to decide whether to hold or sell an asset [5].

Section 4.2.4.2 discusses the importance of integrating the proposed system with big data infrastructure. This is because the data sets recommended in the proposed system are of large volume and must flow to the system in real-time from various data sources. However, due to a lack of real-time data in the prototype, the system was developed with static data, and big data infrastructure was not used. The prototype was developed using the Python programming language and the Plotly Dash package. The reason for utilizing Plotly Dash is that it has user-friendly interactive features and extensive data visualization tools [223]. Another rationale for this decision is the ease of system integration, as modeling is done with Python through TensorFlow and scikit-learn modules. The remainder of this section details the scope of the prototype.

## 5.1 Data Gathering

WTI Futures are traded between 6:00 p.m. and 5:00 p.m. in the United States. Many transactions occur every minute in this market, which can be traded practically 24 hours a day, and the price might fluctuate by the minute or even the second. As a result, real-time data usage is critical for trading decisions. However, access to real-time data for the proposed system's prototype was impossible. This is because this data is only available through commercial data providers. Furthermore, within this study's scope, it is not possible to build an infrastructure for storing and processing such data. Although it is possible to access some data in real time, most of the data used in the system is published at different intervals. For this reason, all data was collected for the most appropriate time interval that can be accessed as open source and artificial data was generated for specific functions. The operations performed are described in detail under the proper headings below.

### 5.1.1 Fundamental Analysis

As explained in Sections 2.1.8 and 4.3.1, the main objective of the Fundamental Analysis component in the proposed system is to forecast the crude oil price using all data sets that affect the crude oil price. Since the prototype targets traders in the WTI Futures market, the WTI Spot price and the WTI Front Month Futures price are predicted daily and weekly. Therefore, studies in the literature on daily and weekly crude oil price forecasts are analyzed, and all data sets published daily are merged and used for modeling. The same data set is used for daily and weekly price forecasting. Table 4 contains information on the dataset used for WTI Spot and WTI Front Month Futures price forecasting.

Table 4:Data Used for WTI Spot and WTI Front Month Futures Price Forecasting

| **Data** | **Studies** | **Group** | **Source** |
|---|---|---|---|
| Gold Price | [1],[80],[84],[85],[88], [158],[159] | Commodity Markets | Investing |
| Natural Gas Price | [156] | Commodity Markets | Investing |
| Exxon Mobil Corporation (XOM) Stock Price | [156] | Commodity Markets | Investing |
| Silver Price | [80] | Commodity Markets | Investing |
| Heating Oil Spot Price | [85], [157] | Commodity Markets | EIA |
| New York Harbor Conventional Gasoline Spot Price | [157] | Commodity Markets | EIA |
| US Gulf Coast Conventional Gasoline Spot Prices | [157] | Commodity Markets | EIA |
| WTI Spot Price/WTI Front Month Futures Price | [157], [173] | Commodity Markets | Investing |
| US Crude Oil Stock (including SPR) | [35] | Commodity Markets | EIA |
| Baltic Exchange Dirty Tanker Index | [35] | Commodity Markets | Investing |
| S&P GSCI Non-Energy Index | [35] | Commodity Markets | S&P Global |
| US 10-Year Bond Yield | [1] | Economic Indicators | FRED |
| US Federal Fund Rate | [157] | Economic Indicators | FRED |
| Three-month U.S. Treasury Bill Rate | [35] | Economic Indicators | FRED |
| US Economic Uncertainty Index (EPU) | [114] | Economic Indicators | FRED |
| US Dollar Index | [1],[85],[88],[114],[158], [178],[226] | Exchange Rates | Investing |
| USD/Euro Exchange Rate | [80] | Exchange Rates | Investing |
| USD/CNY Exchange Rate | [80] | Exchange Rates | Investing |
| USD/Taiwan Dollar Exchange Rate | [84] | Exchange Rates | Investing |
| S&P 500 Index | [1], [35], [80], [85], [157] | Financial Markets | Investing |

Table 4 continued:

| | | | |
|---|---|---|---|
| Dow Jones Industrial Index | [1], [114], [173] | Financial Markets | Investing |
| Bitcoin Price | [80] | Financial Markets | Investing |
| CBOE Volatility Index (VIX) | [157] | Financial Markets | Investing |
| MSCI World | [35] | Financial Markets | Investing |
| EMA, MACD, RSI, MA, Momentum, Williams %R, Stochastic Oscillator, OBV, Volume | [88] | Technical Indicators | Calculated with price and volume data obtained from Investing. |

The daily WTI Spot Price and WTI Front Month Futures price, the target values, were obtained from www.investing.com using Python language and Selenium library via web scraping method. Various academic research has used data from Investing.com [160], [173]. The data covers the period between 22.09.2000 and 17.04.2023, with 5653 price data in total. The data contains daily opening, closing, high and low values, and volume information. The open and close prices reflect the beginning and end of the price in the specified time interval, respectively, while the high and low prices represent the lowest and highest prices in that time interval [212]. Volume, on the other hand, shows the number of transactions in that time.

The data utilized in the modeling is separated into five main groups, as seen in Table 4. The first group is commodity market-related data and includes different commodity prices such as gold, silver, natural gas, heating oil, New York Harbor Conventional Gasoline, and US Gulf Coast Conventional Gasoline. Data on gold, natural gas, and silver prices are also taken from investing.com for the period between 22.09.2000 and 17.04.2023. The U.S. Energy Information Administration (EIA) is responsible for transparently providing energy-related statistics to the public. Heating Oil Spot Price, New York Harbor Conventional Gasoline Spot Price, US Gulf Coast Conventional Gasoline Spot Price, and US Crude Oil Stock (including SPR) data were obtained from www.eia.gov using the same methodology and period. Exxon Mobil Corporation is one of the world's largest oil and gas companies; therefore, its stock price was obtained from investing.com and included in the data set. The Baltic Exchange Dirty Tanker is an index for tracking unrefined oil shipping rates on specific routes. The relevant data was obtained from investing.com, but the starting date of the data is 06.07.2012. Therefore, the missing data was filled using the backward fill method of the Pandas library. The reason behind that is by duplicating the latest accessible value, backward filling keeps the overall trend and pattern of the time series. The S&P GSCI Non-Energy Index is an index that measures the performance of a collection of non-energy commodities. Standard & Poor's (S&P) is a financial company that measures this index. The required data was gathered from www.spglobal.com beginning on 02.01.2009, and the missing data was filled using the backward fill method because by duplicating the latest accessible value, backward filling keeps the overall trend and pattern of the time series.

The Federal Reserve Economic Data (FRED) (www.fred.stlouisfed.org) is the source of all data in the economic indicators group. FRED is a database where various economic and financial data are published by the Federal Reserve Bank of St. Louis and is frequently used in academic studies [37], [38], [160]. All the Exchange Rate and Financial Market group data were gathered from investing.com during the same period and with the same procedures. Only the Bitcoin Price data started on October 26, 2010, and the missing data was filled using the abovementioned process. The technical indicators used in the modeling are calculated using the opening, closing, high, low price, and volume data of WTI Spot and WTI Front Month Futures. It is essential to employ technical indicators in prediction models in fundamental analysis to integrate the advantages of both of these approaches.

### 5.1.2 Technical Analysis

The opening, closing, high and low prices of the WTI Front Month Futures contract, and the volume data are all required for technical analysis. This data was obtained from investing.com daily and weekly between September 22, 2000, and April 17, 2023, as specified in Section 5.1.1. within this study's scope, it was impossible to access this data in real-time, and data from non-commercial data providers are published daily. However, the prototype requires hourly and 4-hour data for technical analysis. To address this issue, hourly and 4-hourly artificial data have been generated utilizing the daily price data's opening, closing, high and low values, and volume. When real-time data is available, technical analysis using data starting at 1 minute may produce superior results for short-term transactions.

### 5.1.3 Sentiment Analysis

The sentiment analysis module examines how daily news flows affect the WTI price. The data used to observe this impact was scraped from oilprice.com using Python's Selenium library. Oilprice.com is a popular website that provides energy market news and articles [2]. Various academic research on sentiment analysis in the crude oil market has used data from this site as a source of data [111], [115], [155]. The dataset contains 11558 crude oil-related news headlines from June 15, 2011, to April 3, 2023. Instead of utilizing the entire news article, only the headlines were used. This is because the headlines provide the most crucial information [2] and contain fewer irrelevant information and words [224].

## 5.2 Modeling and Implementation

The modeling and implementation steps for the prototype's components are discussed below.

### 5.2.1 Fundamental Analysis

The Fundamental Analysis module's primary objective is to forecast crude oil prices using intelligent prediction models based on the factors influencing the oil price. Daily and weekly price forecasts for WTI Spot and WTI Front Month Futures were made in this context utilizing the data sources indicated in Section 5.1.1. This section first describes the steps taken to find the best-performing model and shares the model's results that forecast the WTI Front Month Futures price. Then the interface functionalities that display the model results to the users are detailed.

#### 5.2.1.1 Data Preprocessing

Python was used to aggregate the obtained data to cover the same time interval. Several pieces of data were missing since several markets were closed on weekends and holidays. This data was filled in using the linear interpolation approach. Interpolation is a mathematical method to estimate missing data between known and discrete data [225]. The “interpolate” function of the Python Scipy library was used to perform this operation.

The features in the dataset have very different values, and the difference between the highest and lowest values is relatively high. For modeling, this aspect of the dataset reduces forecast accuracy, and the data should be adjusted to a similar scale [212]. As a result, the Min-Max Normalization was employed in this study. Min-Max normalization uses linear methods to scale data within a specific range [226]. The data was normalized between 0 and 1 using the MinMaxScaler function in the Scikit-learn library's Preprocessing module. The Isolation Forest, an unsupervised machine learning algorithm, was subsequently used to remove the outliers in the dataset [227]. After this process, 283 data were identified as outliers and removed from the data set.

It is crucial to use the appropriate feature set to improve the model's performance. For this reason, feature importance analysis was performed using Random Forest, a machine learning algorithm, and the 30 factors ranked as the most important in predicting the target value were given as input to the prediction model. The results of the feature importance analysis are shown in Figure 7. Past price and technical indicators data are not included in the feature importance analysis as they are directly related to the price.

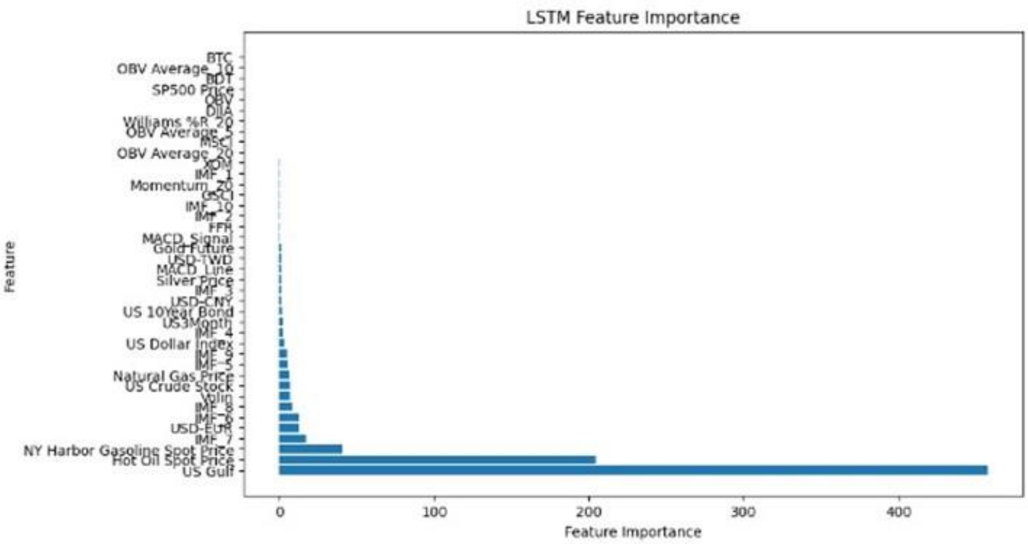


Figure 7:Feature importance analysis results

Since crude oil price data is non-linear and non-stationary, most studies in the literature try to capture the hidden structures in this data by using different decomposition methods [2], [87], [88], [110], [169]. Therefore, to improve the performance of the prediction model in the prototype, ten intrinsic mode functions were identified using Empirical Mode Decomposition (EMD) and fed into the prediction model. EMD is a signal-processing technique used to preprocess data [228]. It breaks down non-stationary time series data into multiple empirical modes to capture hidden patterns and remove noise.

**5.2.1.2 Model Evaluation Metrics**

The Mean Squared Error (MSE) and Mean Absolute Error (MAE) metrics were used to evaluate the model's performance. MSE is the average squared difference between the actual and predicted values, while MAE is the average absolute difference between them. The mathematical formulations of the measurement metrics are as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

where $y_i$ is actual, and $\hat{y}_i$ is predicted values.

#### 5.2.1.3 Modeling and Predictions

This section presents the model architecture for daily forecasting of the WTI Front Month Futures price and the results. The Long Short-Term Memory (LSTM) model is employed as the prediction model due to its superior forecasting performance in time series data and widespread use in the crude oil price forecasting literature, as explained in Section 4.3.1.3.

Hyperparameters play a critical role in the performance of deep learning models; therefore, optimizing hyperparameter parameters is essential [2]. The Grid Search method was utilized for parameter optimization, and the results are presented in Table 5.

Table 5: Hyperparameter Optimization

| **Hyperparameters** | **Values** | **Selected** |
|---|---|---|
| Hidden layers | 1-5 | 3 |
| Batch size | 16,32,64,128,256 | 64 |
| Optimizer | Adam, SGD, Adagrad | Adam |
| Epoch | 50-300 | 200 |
| Learning Rate | 0.1, 0.01, 0.001 | 0.01 |

Forecasting was carried out using a 30-day sliding window after training the model with all the data. This entails forecasting the 31st day's price based on the previous 30 days' data. The data is divided into 80 percent training and 20 percent test data. The training data is used to train the model, and the test data is used to validate the model's performance. After removing outliers, 5369 data points remained from 5652 data points collected between 22.09.2000 and 17.04.2023. Between 02.01.2020 and 17.04.2023, 848 data points were used for testing, while the remaining 4521 data points were used to train the model. For the test set, the average price was $65.24, the standard deviation was $21.48, and the average price variation was $1.39.

Figure 8 depicts a graph of predicted and actual prices based on the model's outcomes. According to the results, the MSE was 1.63, while the MAE was 0.807. These results are based on non-normalized data and mean that the model predicts the WTI Front Month Futures price for the following day with an average deviation of $0.807. It is possible to say that our proposed hybrid EMD-LSTM model, which is trained using data sources identified from the literature, outperforms the results of most studies in the literature on daily crude oil price forecasting. However, such a comparison is not entirely accurate as these studies estimate using data from different time periods. Also, the main purpose of models in the prototype is not to achieve the highest level of performance, but to demonstrate the feasibility of the system. The results of studies in the literature that share the outcomes of non-normalized models in daily crude oil price forecasting are presented in Table 6.

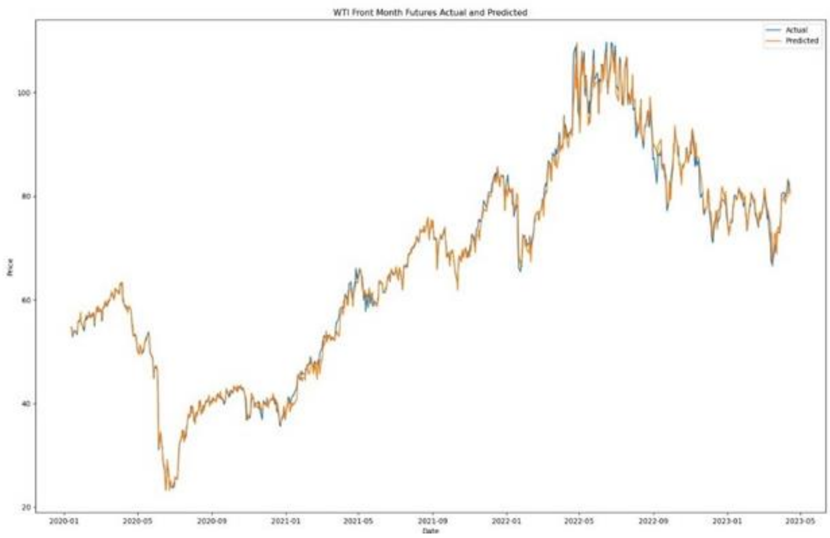


Figure 8:Results of the Prediction Model

Table 6: Model Outcomes of Different Studies

| Study | MAE | Crude Oil | Period |
|---|---|---|---|
| [156] | 13.74 | WTI spot | March 2018-June 2021 |
| [85] | 1.36 | WTI spot | January 2010-December 2013 |
| [84] | 1.19 | WTI spot | May 2009-December 2010 |
| [80] | 1.00 | WTI spot | February 2020- June 2022 |
| [114] | 0.65 | WTI spot | July 2011- July 2021 |
| Our Model | **0.80** | WTI futures | September 2000- April 2023 |

When the predictions of the model in the test set are compared to the actual data, it is observed that the model captures price trend changes with 76 percent accuracy. When only a long position was opened at the points of trend change based on the model's forecasts to run a simple profitability test, it was revealed that 11% profit was produced over the test data period when transaction costs were not considered.

### 5.2.1.4 Implementation

The dataset is categorized into financial markets, technical indicators, commodity markets, exchange rates, and economic indicators. For daily and weekly forecasting, 240 models were trained separately and in various combinations for these five datasets. This allows users to select the data group they want, view the predictions made by the models trained with these data sets, and gain more insight into the drivers of market direction. Figure 9 depicts a screenshot of this functionality. The abovementioned forecasting model, which produces the best results for forecasting the WTI Front Month Futures price, was employed in all forecasting models. Once users have selected the data sets, they want to train the model with, they can view the predicted values of the WTI Spot price and the WTI Front month Futures price for the following day and week.

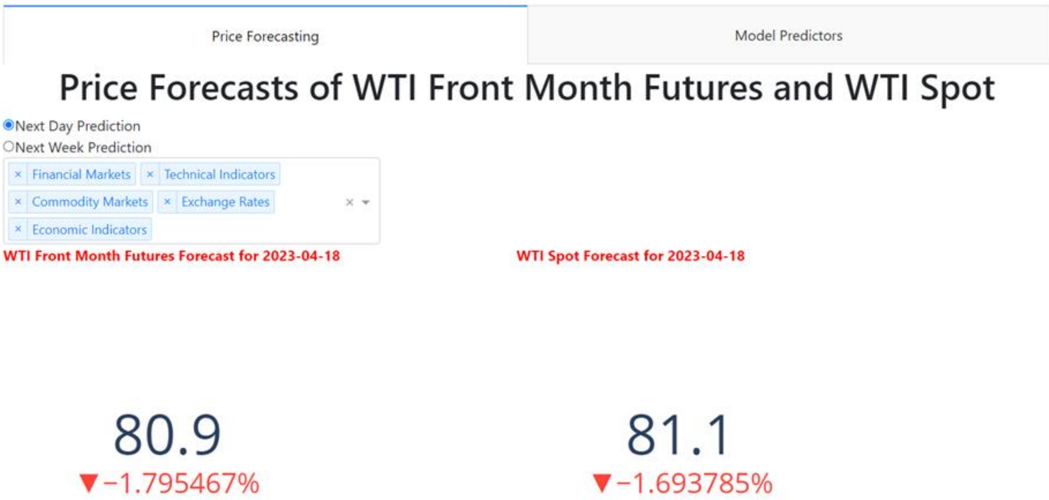


Figure 9:Price forecasting functionality

In addition, the Model Predictors function allows users to view the characteristics of the data groups used in the prediction model and to analyze graphs of historical changes of the different features against each other. A screenshot of this functionality is shown in Figure 10.

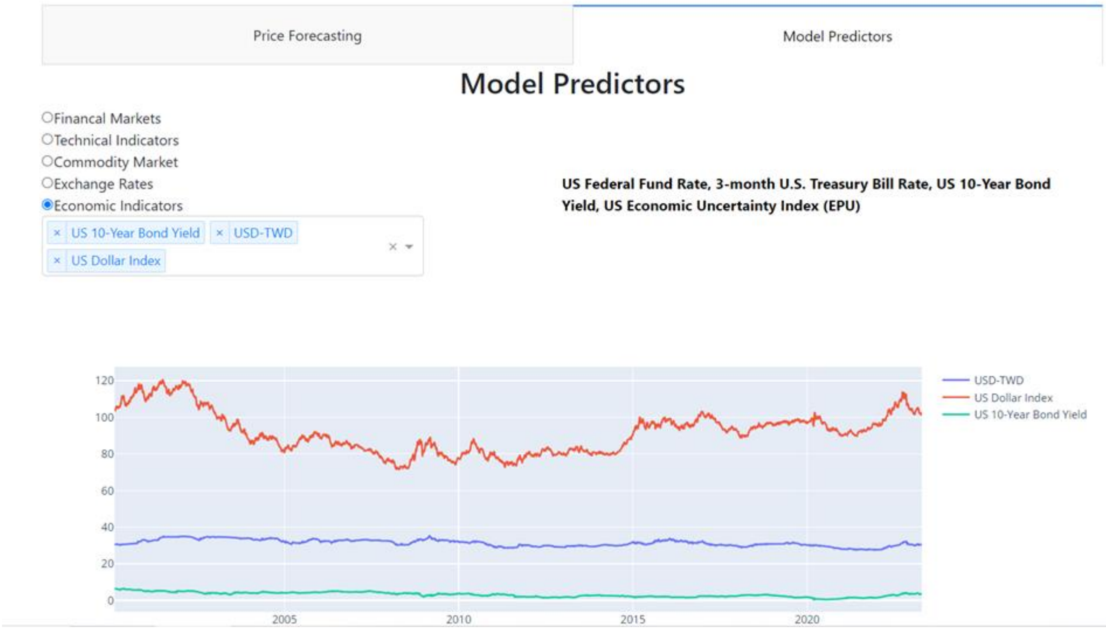


Figure 10:Model predictors in fundamental analysis component

### 5.2.2 Technical Analysis

The technical analysis module runs no models but offers chart analysis, technical indicators, and trading rules for WTI Front Month Futures. Users can analyze charts created with hourly, 4-hourly, daily, and weekly data for candlestick analysis. Screenshots of the hourly and daily candlestick charts are shown in Figure 11.

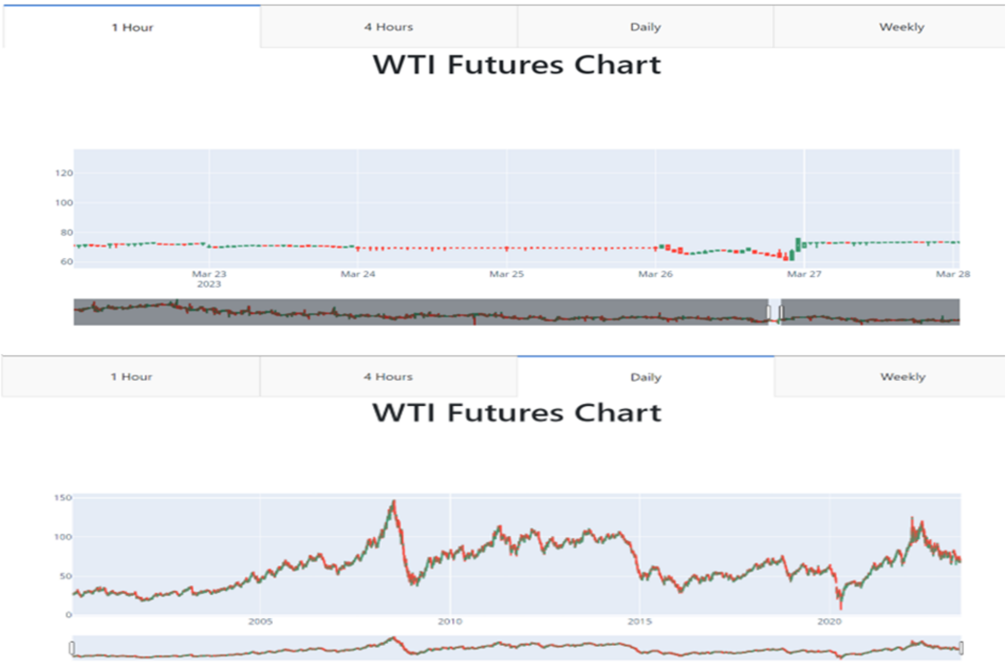


Figure 11:Candlestick page

Users can also compare the volume chart with the candlestick chart for any interval. Figure 12 depicts a screenshot of the 4-hour volume chart and the candlestick chart.


Figure 12:Chart analysis with volume

In addition to chart analysis, users can view the charts of popular technical indicators commonly employed in crude oil trading by calculating them for the desired time intervals. Technical indicators that can be displayed are Moving Average (MA), Relative Strength Index (RSI), Bolinger Bands, Moving Average Convergence Divergence (MACD), Scholastic Oscillator, Fibonacci Retracement, Williams %R, Momentum, Average True Range (ATR), Commodity Channel Index (CCI), Rate of Change (ROC) and On Balance Volume (OBV). One or more technical indicators can be displayed alongside the candlestick chart, helping traders better assess the price movement. Users can view the charts of the technical indicators they have selected on an hourly, 4-hourly, daily, and weekly basis after inputting the parameters required for the technical indicators interactively, they have selected on the screen that will open. Figure 13 depicts a screenshot of the technical indicators and the parameters that should be entered into the system with the selected indicators.

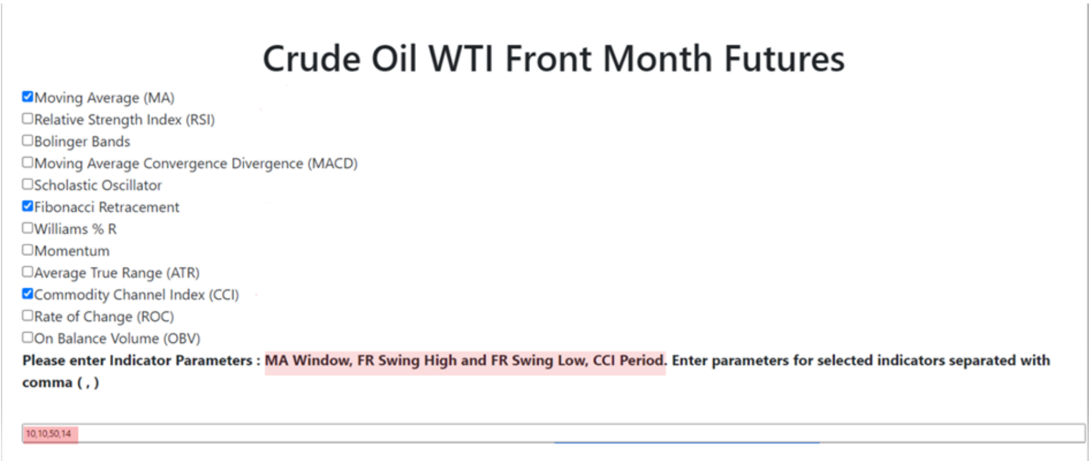


Figure 13:Technical indicators

Figure 14 depicts the charts plotted using the chosen technical indicators and the parameters specified.

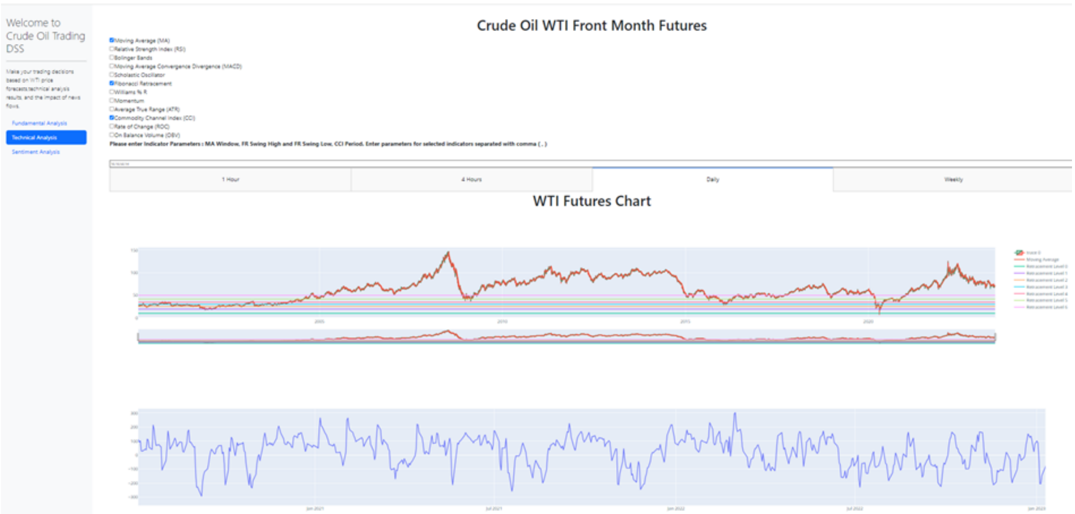


Figure 14:Selected technical indicators with candlestick chart

The technical trading rules are the final element of the technical analysis component. Technical trading rules, like chart analysis and technical indicators, can be calculated hourly, four-hourly, daily, and weekly and generate different signals accordingly. The technical trading rules used are RSI (14), MACD (12,26,9), Williams %R (14, -20, -80), CCI (14), ATR (14), ROC (14), STOCH, STOCHRSI, ADX (14), Ultimate Oscillator, Bull/Bear (13), Highs/Lows (14) and MA (5,10,20,50,100,200). The signals generated by the trading rules and the related buy-sell-neutral signals are shown for the specified interval. Since these trading rules may provide different signals, users can also view a summary signal incorporating all signals. The page of technical trading rules generated for daily data is depicted in Figure 15.

| Technical Trading Rules | Scores | Signal |
|---|---|---|
| RSI (14) | 50.558 | Neutral |
| MACD (12,26,9) | -0.532 | Neutral |
| Williams %R (14,-20,-80) | -60.755 | Neutral |
| CCI (14) | 0.339 | Neutral |
| ATR (14) | 2.424 | Neutral |
| ROC (14) | 0.271 | Neutral |
| STOCH | 39.245 | Buy |
| STOCHRSI | 0.833 | Sell |
| ADX(14) | 13.411 | Neutral |
| Highs / Lows (14) | 8.74 | Buy |
| Ultimate Oscillator | 0.548 | Sell |
| Bull Bear (13) | 1.579 / -0.851 | Sell |
| MA(5) | 71.184 | Sell |
| MA(10) | 70.302 | Buy |
| MA(20) | 70.931 | Sell |
| MA(50) | 72.86 | Sell |
| MA(100) | 74.486 | Sell |
| MA(200) | 77.711 | Sell |

| Positive | Negative | Neutral | Summary |
|---|---|---|---|
| 3 | 8 | 7 | Strong Sell |

Figure 15:Technical trading rules

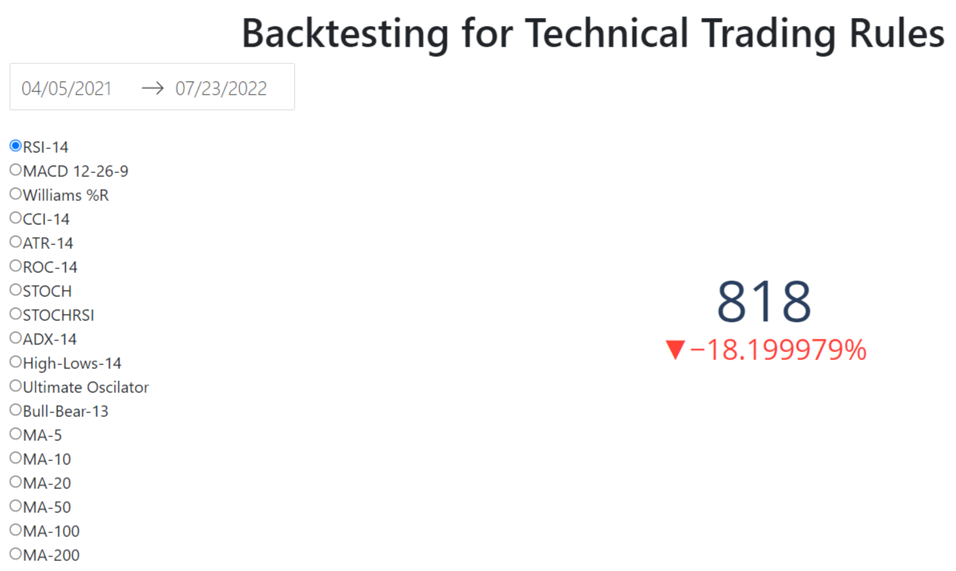


Figure 16: Back testing for technical trading rules

Figure 16 depicts the technical trading rules back test functionality. The profitability of trades conducted with 1000 dollars is shown to users for the date range set by the users and the buy-sell-neutral signals generated by the specified technical trading rule. When the rule produces a buy signal, go long; when the rule gives a sell signal, go short; and when the rule is neutral, no transaction is conducted. As a result of these transactions, users are shown the profit and loss situation once the position is closed. In Figure 16, between 04/05/2021 and 07/23/2022, 1000 dollars were reduced to 818 dollars as a result of trades made with the buy and sell signals generated by the RSI(14) technical trading rule. This means that the signals produced by the RSI(14) rule for the specified time period resulted in an 18% loss.

### 5.2.3 Sentiment Analysis

The Sentiment Analysis module aims to determine the impact of daily news flows on the price of WTI. As indicated in Section 5.1.3, 11558 crude oil-related news headlines from June 15, 2011, to April 3, 2023, from oilprice.com were used. CrudeBert [96], a fine-tuned FinBERT model, performs sentiment analysis. The key reason for selecting this model is that FinBert, trained using financial data, fails to identify some crude oil market characteristics. In contrast, CrudeBert, trained exclusively for the crude oil market, obtains more successful sentiment analysis results in this market [96], [116]. After the news headlines are given as input to the model, the model returns a score between [-1, 1]. As this number approaches 1, it indicates that the related news has an increasing effect on crude oil prices, while as it approaches -1, it indicates that the related news has a decreasing effect on crude oil prices. The sentiment score per news item was obtained using the CrudeBert model on a dataset of 11558 news items. Many methods can calculate the daily sentiment score from the model's news-based scores, such as averaging or summing the scores [2], [88]. To validate the model, daily news scores are averaged, and then daily scores are summed. Figure 17 displays the WTI Front Month Futures Price and cumulative score of the CrudeBert model. The graph shows that the model accurately forecasts the effect of news on WTI Futures price. Since more than just news flows affect the crude oil market, the charts cannot be expected to move precisely in the same direction.

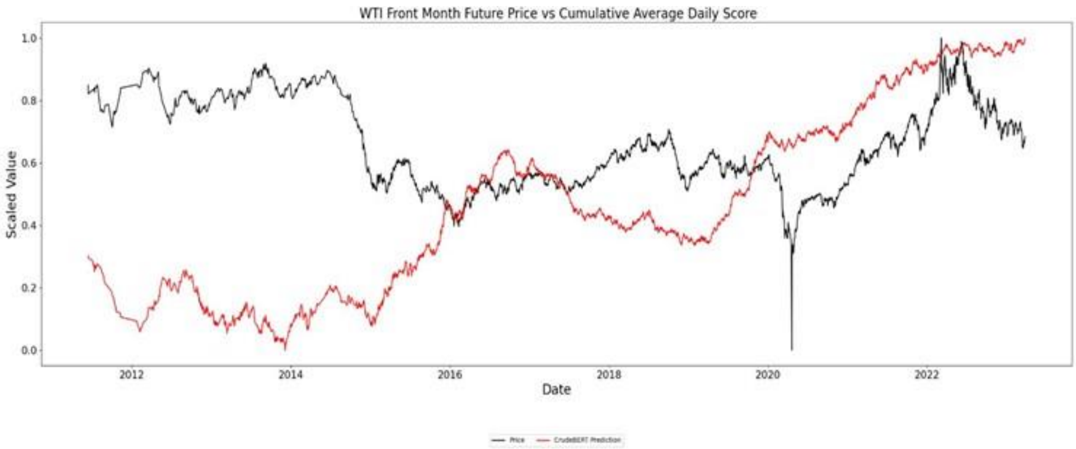


Figure 17:Comparison of WTI Front Month Futures prices and CrudeBert cumulative average daily scores

The model results are presented daily in the prototype to users on a news-based basis. The scores have a threshold of 0.2, which means that news with a sentiment score larger than 0.2 is shown in green (positive), news with a sentiment score less than -0.2 is shown in red (negative), and news between these two values is shown in gray (neutral). Figure 18 depicts news items and scores for a given day. Users can view the news and sentiment scores from any day in the past. Users can also view the cumulative sentiment score calculated for the day of their choice and the sentiment scores calculated using cumulative and exponential average methods over the previous seven days. Figure 19 depicts a screenshot of this functionality.

03/07/2023

| News | Scores |
|---|---|
| Petrobras’ Change In Strategy Provokes Strong Reactions | 0.9156378 |
| Crude Oil Inventories See Weekly Draw As Fuel Inventories Build | -0.97007424 |
| EIA Lowers Forecast For Natural Gas Prices | -0.98565346 |
| Exxon Reconsidering Its Role In Europe Thanks To Windfall Taxes | 0.79283786 |
| U.S. Oil Production To Grow Just 500,000 Bpd This Year | -0.9949789 |
| Output At Huge Kazakh Oilfield Dips On Unscheduled Maintenance | 0.99276745 |
| Russia Has Started Exporting Diesel To Saudi Arabia | -0.9913808 |
| The Kremlin Will Not Recognize Any Price Cap On Russia’s Oil | -0.031770438 |

Figure 18:Daily news and sentiment scores

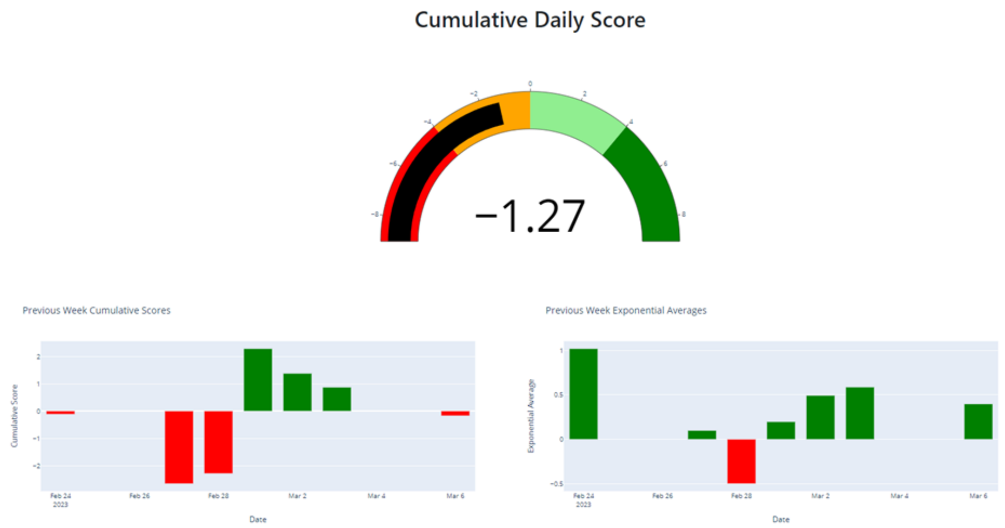


Figure 19:Daily and weekly sentiment scores

70

# CHAPTER 6

# DISSUSSION AND CONCLUSION

## 6.1 Dissussion

The main objective of the proposed decision support system is to support the decision-making processes of individual traders who trade crude oil for profit by bringing together a wide variety of features and different analysis methods affecting the oil market. In line with this objective, the advantages of the proposed system and the implementation are presented as follows.

The factors influencing crude oil prices vary significantly, making price forecasting challenging [1]. These factors include economic considerations, political instability, commodity pricing, and pandemics like Covid 19 that affect the globalized world [1]. It is quite difficult for individual traders to keep track of all these factors, and this can lead to information asymmetry in the market. In this context, the first advantage of the proposed system is that it identifies all factors that have the potential to affect the crude oil price through a systematic literature review and proposes an infrastructure to bring together the data published by several different sources. Furthermore, these features are classified into several categories, giving users the flexibility to choose which data sets to employ in the forecasting model. By comparing alternative forecasts created with data sets thought to be suitable for the specific trading technique, it is possible to better assess which factors have an impact on the market direction and to what extent.

In order to conduct profitable trading activities, the price needs to be accurately forecasted after determining the factors affecting the market. Fundamental analysis and technical analysis are the two fundamental price forecasting approaches employed to forecast the price of any financial instrument. As explained throughout the study, these two strategies produce considerable outcomes on their own, but combining them produces significantly more meaningful and profitable trading activities [229]. The majority of research in the literature indicate that both of these methods should be utilized in combination [131], [229]–[231]. As a result, both fundamental and technical analysis literature were examined in detail, and various approaches were integrated into the proposed system. Generally, technical analysis is used to anticipate short-term price movements, whereas fundamental analysis is used to predict long-term price movements [144]. As a result, the proposed approach is flexible enough to enable trading strategies that involve time periods ranging from intraday to yearly.

The ability to accurately predict price and trend is critical for trading decision support systems so that suggested trading positions can achieve profitable outcomes [232], [233]. Therefore, in the proposed system, the factors affecting the crude oil price are used for fundamental analysis and price prediction is made with machine learning models. It is possible to make intraday, daily, weekly, monthly and annual price forecasts according to the trading strategy used by selecting the appropriate data sets from the presented data set. In the prototype, the data sets that can be used for daily and weekly price prediction were determined and predictive models was implemented for the next day and the next week price prediction using different data groups according to the selections of the users. The systematic literature review also provided intelligent methods to predict crude oil price over various time intervals. The proposed system is sufficiently flexible to use the most accurate models as the baseline after training various models with updated datasets.

When making a trading decision, traders should also consider the relationship between the predictors used in the price forecasting model and the target value. As a result, another advantage of the system is the ability to create charts where users can examine the relationship between the historical changes of the model predictors and the crude oil price.

Chart analysis, technical indicators, and technical trading rules are all part of technical analysis. The proposed decision support system also has the advantage of including different technical analysis approaches. Traders can examine historical candlestick and volume charts of WTI prices for various time intervals based on their trading strategy. Furthermore, users can interactively display the technical indicators often used in the crude oil literature for the time intervals they prefer, in accordance with their trading strategies which may be at low or high frequency. The profitability of technical trading rules is debated in academic literature, as stated in Section 2.1.10. Therefore, the proposed system's technical trading rules can be back tested for profitability utilizing historical data at time intervals interactively selected by the users. By comparing periods in which the price has similar tendencies, the rules appropriate for the selected trading strategy can be defined more accurately.

Sentiment analysis is another crucial analysis method that should be considered when making a trading decision. Another advantage of the proposed system is that it provides a tool for users to monitor the impact of daily news flows on crude oil prices. In addition to news-based scores, users can also view the daily cumulative score, which provides greater insight into the general sentiment affecting the market and, as a result, the market's direction. The impact of news on the price of oil is continuous in the oil market, which means that the price is impacted by both current day and prior day news [2]. For this reason, users were also presented with the scores of the previous weeks calculated by cumulative and exponential average methods. In addition, the proposed system is flexible enough to incorporate not only the impact of news feeds on crude oil price but also investor sentiment analysis that can be measured using social media posts or search engine logs.

To summarize, the proposed system identifies all of the factors that can affect the oil market, provides the necessary infrastructure to aggregate and process these data, and brings together various analysis methods that can be used to support trading decisions from a holistic perspective, giving the user flexibility. In this regard, the study differs from others in the literature, and to the best of our knowledge, it is the first decision support system study that brings all these concepts together for the crude oil market.

## 6.2 Summary

This study aims to propose an intelligent decision support system for the crude oil trading problem. First, the literature investigated the details of the crude oil market and the analysis methods utilized to profit in this and other financial markets. Then the decision support literature was reviewed, and it discovered that no study gives a comprehensive approach traders can use to profit in the crude oil market.

A systematic literature review is conducted to identify and present all factors that can impact crude oil prices. The methods of technical analysis, fundamental analysis, and sentiment analysis that can be used to determine the crude oil price and market direction have been extensively explored, and the technology, tools, and procedures that can be utilized to undertake these analyses have been discovered. Afterward, a comprehensive decision support architecture is proposed that consists of real-time data collection and integration, modeling, and intelligence and visualization layers. Crude oil is traded in the spot and derivatives markets via various futures, spot, and options contracts. The most crucial thing to profit from this market is anticipating price volatility while keeping watch of all market trends. Individual traders may find it challenging to get the necessary information to track price fluctuations in this market because a wide range of factors affect the price of crude oil. The proposed decision support system's main objective is eliminating this information asymmetry for individual traders.

Finally, a prototype of the proposed system was created. The crude oil price was forecasted using data from the literature in this prototype, and the results were better than most of the research in the literature. Furthermore, news feeds were used, and their impact on the crude oil price was discovered. Validation of the proposed system is based on the realization of the proposed system with the prototype and the models' performances.

## 6.3 Contributions

As highlighted in the literature review, no financial decision support system study covers all analysis methods for a particular financial instrument. As a result, employing intelligent techniques, this study presents a holistic perspective that integrates fundamental, technical, and sentiment analysis methods. When other studies in the literature are investigated, it is discovered that the studies offer a system to help traders' decision-making process for the stock market by employing only technical analysis, fundamental analysis, sentiment analysis, or a combination of two of these. However, in the crude oil literature, various strategies are offered to traders only through price forecasting. In this study, however, firstly, a systematic literature review is conducted to identify all data sources that can be used for crude oil price forecasting. Then, big data infrastructure is proposed to collect these data sources and news flows published in multiple formats from various sources, preprocess them, and make them suitable for modeling. Then, machine learning algorithms for price forecasting and sentiment analysis are investigated and presented. In addition to modeling, technical analysis methods such as chart analysis, technical indicators, and technical trading rules identified in academic and grey literature and used by traders to better interpret short-term price movements have been defined and incorporated into the proposed system. In this regard, unlike other studies in literature, the proposed approach combines the various data types and analysis techniques that can be used for crude oil trading with the appropriate infrastructure.

A prototype is developed as well to illustrate the proposed system's applicability and feasibility. Daily and weekly forecasts of WTI Spot and WTI Front Month Futures prices were performed using real data in this prototype. It is possible to say that the modeling findings performed well, outperforming the majority of the studies studied in the literature. Furthermore, users are given the option to back-test signals generated by technical trading rules using historical data.

## 6.4 Limitations and Future Work

The proposed system only provides recommendations based on the price and trend forecast of crude oil price, effects of news on the price, and technical analysis. Because access to real-time data was not possible during the prototype's design, the proposed big data infrastructure could not be implemented, and for some functions, artificial data was utilized instead of real data.

Another limitation of the proposed system is related to sentiment analysis. The dataset used for sentiment analysis in the prototype contains only news related to crude oil. Using such a data set makes it difficult to measure the impact of exogenous factors such as Covid-19, which significantly affect the oil market. For sentiment analysis to perform as proposed in future studies, the impact of financial and non-financial news should be measured using distinct models. Furthermore, it is critical that the news in the data source be published at the time of the news release to measure the news's impact at the right moment. In addition to news sources, sentiment analysis with social media posts or search engine logs might be included to assess investors' perceptions of the market.

A systematic literature review was used to identify the data that can be utilized in the proposed system. However, new factors affecting the oil price may arise over time. The proposed system does not include an approach for updating the data sets used for modeling.

Furthermore, the prototype was only tested on the WTI Futures market. It would be more beneficial in future studies to establish a system that incorporates relevant analyses for all other main oil types, such as Brent, Dubai, and others.

The proposed system can be used as a part of a broader algorithmic trading system in future studies. Instead of generic technical trading rules, it may be possible to generate new rules from historical data through testing the profitability.

## REFERENCES


[1] S. Urolagin, N. Sharma, and T. K. Datta, "A combined architecture of multivariate LSTM with Mahalanobis and Z-Score transformations for oil price forecasting," *Energy*, vol. 231, p. 120963, Sep. 2021, doi: 10.1016/j.energy.2021.120963.

[2] H. Jiang, W. Hu, L. Xiao, and Y. Dong, "A decomposition ensemble based deep learning approach for crude oil price forecasting," *Resources Policy*, vol. 78, p. 102855, Sep. 2022, doi: 10.1016/j.resourpol.2022.102855.

[3] "IMF Working Papers Volume 2008 Issue 006: Financial Instruments to Hedge Commodity Price Risk for Developing Countries (2008)," *imfsg*. https://www.elibrary.imf.org/view/journals/001/2008/006/001.2008.issue-006-en.xml (accessed Jul. 11, 2023).

[4] K. Movagharnejad, B. Mehdizadeh, M. Banihashemi, and M. S. Kordkheili, "Forecasting the differences between various commercial oil prices in the Persian Gulf region by neural network," *Energy*, vol. 36, no. 7, pp. 3979–3984, Jul. 2011, doi: 10.1016/j.energy.2011.05.004.

[5] S. Deng *et al.*, "Prediction and Trading in Crude Oil Markets Using Multi-Class Classification and Multi-Objective Optimization," *IEEE Access*, vol. 7, pp. 182860–182872, 2019, doi: 10.1109/ACCESS.2019.2960379.

[6] S. Rath, P. Samal, and J. Behera, "Fundamental and Technical Analysis in Future Trading," vol. 2, pp. 60–63, Apr. 2020.

[7] J. J. Murphy, *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. Penguin, 1999.

[8] X. Deng, V. Bashlovkina, F. Han, S. Baumgartner, and M. Bendersky, "What do LLMs Know about Financial Markets? A Case Study on Reddit Market Sentiment Analysis." arXiv, Dec. 21, 2022. doi: 10.48550/arXiv.2212.11311.

[9] S. H. Irwin and D. R. Sanders, "Financialization and Structural Change in Commodity Futures Markets," *Journal of Agricultural and Applied Economics*, vol. 44, no. 3, 2012, doi: 10.1017/s1074070800000481.

[10] D. Wang and T. Fang, "Forecasting Crude Oil Prices with a WT-FNN Model," *Energies*, vol. 15, no. 6, 2022, doi: 10.3390/en15061955.

[11] "NYMEX - CME Group." https://www.cmegroup.com/content/cmegroup/en/company/nymex.html (accessed Jul. 10, 2023).

[12] "CME Exchange - CME Group." https://www.cmegroup.com/content/cmegroup/en/company/cme.html (accessed Jul. 10, 2023).

[13] N. Blasco, R. Santamaria, and N. Satrústegui, "The witching week of herding on bitcoin exchanges," *Financial Innovation*, vol. 8, no. 1, 2022, doi: 10.1186/s40854-021-00323-4.

[14] J. Malhotra and A. Corelli, "The Relative Informativeness of Regular and E-Mini Euro/Dollar Futures Contracts and the Role of Trader Types," *Risks*, vol. 9, no. 6, 2021, doi: 10.3390/risks9060111.

[15] M. S. Tessmann, C. E. Carrasco-Gutierrez, and A. V. Lima, "Determinants of Corn and Soybean Futures Prices Traded on the Brazilian Stock Exchange: An ARDL Approach," *International Journal of Economics and Finance*, vol. 15, no. 1, 2022, doi: 10.5539/ijef.v15n1p65.

[16] "ICE Futures U.S." https://www.theice.com/futures-us (accessed Jul. 10, 2023).

[17] L. Yu, S. Wang, and K. K. Lai, "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm," *Energy Economics*, vol. 30, no. 5, pp. 2623–2635, Sep. 2008, doi: 10.1016/j.eneco.2008.05.003.

[18] B. Algieri, "A roller coaster ride: an empirical investigation of the main drivers of the international wheat price," *Agricultural Economics*, vol. 45, no. 4, pp. 459–475, 2014, doi: 10.1111/agec.12099.

[19] L. Kilian and D. P. Murphy, "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil," *Journal of Applied Econometrics*, vol. 29, no. 3, pp. 454–478, 2014, doi: 10.1002/jae.2322.

[20] A. Cologni, E. Scarpa, and F. Sitzia, "Big Fish: Oil Markets and Speculation." Rochester, NY, Jun. 08, 2015. doi: 10.2139/ssrn.2615868.

[21] C.-L. Chang, M. McAleer, and R. Tansuchat, "Crude oil hedging strategies using dynamic multivariate GARCH," *Energy Economics*, vol. 33, no. 5, pp. 912–923, Sep. 2011, doi: 10.1016/j.eneco.2011.01.009.

[22] D. Valenti, "Modelling the Global Price of Oil: Is there any Role for the Oil Futures-spot Spread?," *The Energy Journal*, vol. 43, no. 2, 2022, doi: 10.5547/01956574.43.2.dval.

[23] F. Massacci, C. N. Ngo, J. Nie, D. Venturi, and J. Williams, "FuturesMEX: Secure, Distributed Futures Market Exchange," in *2018 IEEE Symposium on Security and Privacy (SP)*, May 2018, pp. 335–353. doi: 10.1109/SP.2018.00028.

[24] A. H. Amadi, O. F. Ebube, and S. I. Aire, "Effects of Covid-19 on Crude Oil Price and Future Forecast Using a Model Application and Machine Learning," *European Journal of Engineering and Technology Research*, vol. 5, no. 12, 2020, doi: 10.24018/ejeng.2020.5.12.2232.

[25] M. Chiu, "Derivatives markets, products and participants: an overview," Bank for International Settlements, IFC Bulletins chapters, 2012. Accessed: Jul. 11, 2023. [Online]. Available: https://econpapers.repec.org/bookchap/bisbisifc/35-01.htm

[26] A. Andrews, R. Pirog, and M. F. Sherlock, "The U.S. oil refining industry: Background in changing markets and fuel policies," pp. 1–50, Jan. 2011.

[27] R. Su, J. Du, F. Shahzad, and X. Long, "Unveiling the Effect of Mean and Volatility Spillover between the United States Economic Policy Uncertainty and WTI Crude Oil Price," *Sustainability*, vol. 12, no. 16, Art. no. 16, Jan. 2020, doi: 10.3390/su12166662.

[28] P. Maia, "Understanding-Crude-Oil-and-Product-Markets-Primer-High - Trader," *Passei Direto*. https://www.passeidireto.com/arquivo/101668347/understanding-crude-oil-and-product-markets-primer-high (accessed Jul. 11, 2023).

[29] Y. Wang, L. Liu, X. Diao, and C. Wu, "Forecasting the real prices of crude oil under economic and statistical constraints," *Energy Economics*, vol. 51, pp. 599–608, Sep. 2015, doi: 10.1016/j.eneco.2015.09.003.

[30] J. Considine, P. Galkin, and A. Aldayel, "Global Crude Oil Storage Index: A New Benchmark for Energy Policy," *Methodology Papers*, Art. no. ks--2022-mp01, Sep. 2021, Accessed: Aug. 19, 2023. [Online]. Available: https://ideas.repec.org//p/prc/mpaper/ks--2022-mp01.html

[31] Z. Li and H. Zhao, "Not all demand oil shocks are alike: disentangling demand oil shocks in the crude oil market," *Journal of Chinese Economic and Foreign Trade Studies*, vol. 4, no. 1, pp. 28–44, Jan. 2011, doi: 10.1108/17544401111106798.

[32] J. Beckman and G. Nigatu, "Do political factors influence US crude oil imports?," *International Journal of Energy Economics and Policy : IJEEP*, vol. 11, no. 4, 2021.

[33] S. Deng and A. Sakurai, "Crude Oil Spot Price Forecasting Based on Multiple Crude Oil Markets and Timeframes," *Energies*, vol. 7, no. 5, Art. no. 5, May 2014, doi: 10.3390/en7052761.

[34] A. S. Kalaitzi and T. W. Chamberlain, “Manufactured exports, disaggregated imports and economic growth: the case of Kuwait,” *Econ Change Restruct*, vol. 56, no. 2, pp. 919–940, Apr. 2023, doi: 10.1007/s10644-022-09444-x.

[35] H. Miao, S. Ramchander, T. Wang, and D. Yang, “Influential factors in crude oil price forecasting,” *Energy Economics*, vol. 68, pp. 77–88, Oct. 2017, doi: 10.1016/j.eneco.2017.09.010.

[36] J. Zhao, “Exploring the influence of the main factors on the crude oil price volatility: An analysis based on GARCH-MIDAS model with Lasso approach,” *Resources Policy*, vol. 79, p. 103031, Dec. 2022, doi: 10.1016/j.resourpol.2022.103031.

[37] X. Cheng, P. Wu, S. S. Liao, and X. Wang, “An integrated model for crude oil forecasting: Causality assessment and technical efficiency,” *Energy Economics*, vol. 117, p. 106467, Jan. 2023, doi: 10.1016/j.eneco.2022.106467.

[38] B. Lucey and B. Ren, “Does news tone help forecast oil?,” *Economic Modelling*, vol. 104, p. 105635, Nov. 2021, doi: 10.1016/j.econmod.2021.105635.

[39] Y. Liu, Y. Wei, Y. Liu, and W. Li, “Forecasting Oil Price by Hierarchical Shrinkage in Dynamic Parameter Models,” *Discrete Dynamics in Nature and Society*, vol. 2020, p. e6640180, Dec. 2020, doi: 10.1155/2020/6640180.

[40] Y. Wei, J. Liu, X. Lai, and Y. Hu, “Which determinant is the most informative in forecasting crude oil market volatility: Fundamental, speculation, or uncertainty?,” *Energy Economics*, vol. 68, pp. 141–150, Oct. 2017, doi: 10.1016/j.eneco.2017.09.016.

[41] Q. Lu, S. Sun, H. Duan, and S. Wang, “Analysis and forecasting of crude oil price based on the variable selection-LSTM integrated model,” *Energy Informatics*, vol. 4, p. 47, Sep. 2021, doi: 10.1186/s42162-021-00166-4.

[42] J. Melero, J. Iglesias, and A. García, “Biomass as renewable feedstock in standard refinery units. Feasibility, opportunities and challenges,” *Energy Environ. Sci.*, vol. 5, pp. 7393–7420, May 2012, doi: 10.1039/C2EE21231E.

[43] N. Al-Yousef, “Fundamentals and Oil Price Behaviour: New Evidence from Co-integration Tests with Structural Breaks and Granger Causality Tests,” *Australian Economic Papers*, vol. 57, no. 1, pp. 1–18, 2018, doi: 10.1111/1467-8454.12098.

[44] N. H. A. Razek and N. M. Michieka, “OPEC and non-OPEC production, global demand, and the financialization of oil,” *Research in International Business and Finance*, vol. 50, pp. 201–225, Dec. 2019, doi: 10.1016/j.ribaf.2019.05.009.

[45] R. K. Kaufmann, S. Dees, P. Karadeloglou, and M. Sanchez, “Does OPEC Matter? An Econometric Analysis of Oil Prices,” *EJ*, vol. 25, no. 4, Oct. 2004, doi: 10.5547/ISSN0195-6574-EJ-Vol25-No4-4.

[46] O. Elsalih, K. Sertoglu, and M. Besim, “Determinants of comparative advantage of crude oil production: Evidence from OPEC and non-OPEC countries,” *International Journal of Finance & Economics*, vol. 26, no. 3, pp. 3972–3983, 2021, doi: 10.1002/ijfe.1999.

[47] Y. Zou and K. He, “Forecasting Crude Oil Risk Using a Multivariate Multiscale Convolutional Neural Network Model,” *Mathematics*, vol. 10, no. 14, Art. no. 14, Jan. 2022, doi: 10.3390/math10142413.

[48] J. A. Frankel, “The Effect of Monetary Policy on Real Commodity Prices.” in Working Paper Series. National Bureau of Economic Research, Dec. 2006. doi: 10.3386/w12713.

[49] L. Yan, Y. Zhu, and H. Wang, “Selection of Machine Learning Models for Oil Price Forecasting: Based on the Dual Attributes of Oil,” *Discrete Dynamics in Nature and Society*, vol. 2021, p. e1566093, Oct. 2021, doi: 10.1155/2021/1566093.

[50] L. Yan, “Analysis of the International Oil Price Fluctuations and Its Influencing Factors,” vol. 2012, Apr. 2012, doi: 10.4236/ajibm.2012.22006.

[51] Y. Wang and C. Wu, “Energy prices and exchange rates of the U.S. dollar: Further evidence from linear and nonlinear causality analysis,” *Economic Modelling*, vol. 29, no. 6, pp. 2289–2297, Nov. 2012, doi: 10.1016/j.econmod.2012.07.005.

[52] P. Sadorsky, “The empirical relationship between energy futures prices and exchange rates,” *Energy Economics*, vol. 22, no. 2, pp. 253–266, Apr. 2000, doi: 10.1016/S0140-9883(99)00027-4.

[53] W. Mensi, S. Hammoudeh, and S. H. Kang, “Precious metals, cereal, oil and stock market linkages and portfolio risk management: Evidence from Saudi Arabia,” *Economic Modelling*, vol. 51, pp. 340–358, Dec. 2015, doi: 10.1016/j.econmod.2015.08.005.

[54] M. E. Bildirici and C. Turkmen, “Nonlinear causality between oil and precious metals,” *Resources Policy*, vol. 46, pp. 202–211, Dec. 2015, doi: 10.1016/j.resourpol.2015.09.002.

[55] Y.-J. Zhang and Y.-F. Sun, “The dynamic volatility spillover between European carbon trading market and fossil energy market,” *Journal of Cleaner Production*, vol. 112, pp. 2654–2663, Jan. 2016, doi: 10.1016/j.jclepro.2015.09.118.

[56] V. Costantini, F. Gracceva, A. Markandya, and G. Vicini, “Security of energy supply: Comparing scenarios from a European perspective,” *Energy Policy*, vol. 35, no. 1, pp. 210–226, Jan. 2007, doi: 10.1016/j.enpol.2005.11.002.

[57] M. A. Ruiz Estrada, D. Park, M. Tahir, and A. Khan, “Simulations of US-Iran war and its impact on global oil price behavior,” *Borsa Istanbul Review*, vol. 20, no. 1, pp. 1–12, Mar. 2020, doi: 10.1016/j.bir.2019.11.002.

[58] S. Huang, H. An, S. Wen, and F. An, “Revisiting driving factors of oil price shocks across time scales,” *Energy*, vol. 139, pp. 617–629, Nov. 2017, doi: 10.1016/j.energy.2017.07.158.

[59] K. Drachal, “Forecasting the Crude Oil Spot Price with Bayesian Symbolic Regression,” *Energies*, vol. 16, no. 1, Art. no. 1, Jan. 2023, doi: 10.3390/en16010004.

[60] A. B. R. Costa, P. C. G. Ferreira, W. P. Gaglianone, O. T. C. Guillén, J. V. Issler, and Y. Lin, “387-Machine learning and oil price point and density forecasting,” *Energy Economics*, vol. 102, p. 105494, Oct. 2021, doi: 10.1016/j.eneco.2021.105494.

[61] D. Caldara and M. Iacoviello, “Measuring Geopolitical Risk,” Mar. 2022, Accessed: Jul. 12, 2023. [Online]. Available: https://www.federalreserve.gov/econres/ifdp/measuring-geopolitical-risk.htm

[62] G. A. Fontanills, *Getting Started in Commodities*. John Wiley & Sons, 2007.

[63] J. Kaur and K. Dharni, “Data mining–based stock price prediction using hybridization of technical and fundamental analysis,” *Data Technologies and Applications*, vol. ahead-of-print, no. ahead-of-print, Jan. 2023, doi: 10.1108/DTA-04-2022-0142.

[64] W. Dbouk and I. Jamali, “Predicting daily oil prices: Linear and non-linear models,” *Research in International Business and Finance*, vol. 46, pp. 149–165, Dec. 2018, doi: 10.1016/j.ribaf.2018.01.003.

[65] J. Liu and X. Huang, “Forecasting Crude Oil Price Using Event Extraction,” *IEEE Access*, vol. 9, pp. 149067–149076, 2021, doi: 10.1109/ACCESS.2021.3124802.

[66] M. Gülerce and G. Unal, “FORECASTING OF OIL AND AGRICULTURAL COMMODITY PRICES: VARMA VERSUS ARMA,” *Annals of Financial Economics*, vol. 12, p. 1750012, Sep. 2017, doi: 10.1142/S2010495217500129.

[67] M. Aamir, A. Shabri, and M. Ishaq, “Crude oil price forecasting by CEEMDAN based hybrid model of ARIMA and Kalman filter,” *Jurnal Teknologi*, vol. 80, pp. 67–79, Jul. 2018, doi: 10.11113/jt.v80.10852.

[68] L. Quan, "Daqing crude oil price forecast based on the ARIMA model," *The Open Petroleum Engineering Journal*, vol. 8, pp. 457–462, Oct. 2015, doi: 10.2174/1874834101508010457.

[69] M. Aamir, A. Shabri, and M. Ishaq, "Improving forecasting accuracy of crude oil prices using decomposition ensemble model with reconstruction of IMFs based on ARIMA model," vol. 10, pp. 471–483, Dec. 2018, doi: 10.11113/mjfas.v14n4.1013.

[70] S. R. Yaziz, M. Ahmad, L. Nian, and N. Muhammad, "A Comparative Study on Box-Jenkins and Garch Models in Forecasting Crude Oil Prices," *Journal of Applied Sciences*, vol. 11, Jul. 2011, doi: 10.3923/jas.2011.1129.1135.

[71] M. Haque and A. Rahman, "Predicting Crude Oil Prices During a Pandemic: A Comparison of Arima and Garch Models," *Montenegrin Journal of Economics*, vol. 17, pp. 197–207, Jan. 2021, doi: 10.14254/1800-5845/2021.17-1.15.

[72] L. Lin, Y. Jiang, H. Xiao, and Z. Zhou, "Crude oil price forecasting based on a novel hybrid long memory GARCH-M and wavelet analysis model," *Physica A: Statistical Mechanics and its Applications*, vol. 543, p. 123532, Apr. 2020, doi: 10.1016/j.physa.2019.123532.

[73] K. Drachal, "Forecasting crude oil real prices with averaging time-varying VAR models," *Resources Policy*, vol. 74, p. 102244, Dec. 2021, doi: 10.1016/j.resourpol.2021.102244.

[74] Y. Wang, L. Liu, and C. Wu, "Forecasting the real prices of crude oil using forecast combinations over time-varying parameter models," *Energy Economics*, vol. 66, pp. 337–348, Aug. 2017, doi: 10.1016/j.eneco.2017.07.007.

[75] J. Chai, Y. Wang, S. Wang, and Y. Wang, "A decomposition–integration model with dynamic fuzzy reconstruction for crude oil price prediction and the implications for sustainable development," *Journal of Cleaner Production*, vol. 229, pp. 775–786, Aug. 2019, doi: 10.1016/j.jclepro.2019.04.393.

[76] H. Xie, X. Zhang, and S. Wang, "THE MORE THE BETTER: FORECASTING OIL PRICE WITH DECOMPOSITION-BASED VECTOR AUTOREGRESSIVE MODEL," *International Journal of Energy and Statistics*, vol. 01, Apr. 2013, doi: 10.1142/S233568041350004X.

[77] R. Ahmed and A. Shabri, "Fitting GARCH models to crude oil spot price data," *Life Science Journal*, vol. 10, pp. 654–661, Jan. 2013.

[78] D. Xu, T. Chen, and W. Xu, "89-A support vector machine-based ensemble prediction for crude oil price with VECM and STEPMRS," *International Journal of Global Energy Issues*, vol. 38, Jan. 2015, doi: 10.1504/IJGEI.2015.069488.

[79] J. An, A. Mikhaylov, and N. Moiseev, “Oil Price Predictors: Machine Learning Approach,” *International Journal of Energy Economics and Policy*, vol. 9, no. 5, Art. no. 5, Jul. 2019.

[80] X. Gao, J. Wang, and L. Yang, “An Explainable Machine Learning Framework for Forecasting Crude Oil Price during the COVID-19 Pandemic,” *Axioms*, vol. 11, no. 8, Art. no. 8, Aug. 2022, doi: 10.3390/axioms11080374.

[81] Y. Lin, D. Han, J. Du, and G. Jia, “The Mechanism of Google Trends Affecting Crude Oil Price Forecasting,” *SN COMPUT. SCI.*, vol. 3, no. 4, p. 294, May 2022, doi: 10.1007/s42979-022-01195-w.

[82] P. Sohrabi, H. Dehghani, and R. Rafie, “Forecasting of WTI crude oil using combined ANN-Whale optimization algorithm,” *Energy Sources, Part B: Economics, Planning, and Policy*, vol. 17, no. 1, p. 2083728, Dec. 2022, doi: 10.1080/15567249.2022.2083728.

[83] L. Yu *et al.*, “Forecasting crude oil spot price by wavelet neural networks using OECD petroleum inventory levels,” *New Mathematics and Natural Computation (NMNC)*, vol. 07, pp. 281–297, May 2011, doi: 10.1142/S1793005711001937.

[84] S.-C. Huang and C.-F. Wu, “Energy Commodity Price Forecasting with Deep Multiple Kernel Learning,” *Energies*, vol. 11, no. 11, Art. no. 11, Nov. 2018, doi: 10.3390/en11113029.

[85] S. Nanda, C. Aloui, and M. Hamdi, “Comparing Functional Link Artificial Neural Network And Multilayer Feedforward Neural Network Model To Forecast Crude Oil Prices, Economics Bulletin, 2016, vol. 36, issue 4, pages 2430-2442,” *Economics Bulletin*, vol. 36, pp. 2430–2442, Dec. 2016.

[86] T. Mingming and Z. Jinliang, “A multiple adaptive wavelet recurrent neural network model to analyze crude oil prices,” *Journal of Economics and Business*, vol. 64, no. 4, pp. 275–286, Jul. 2012, doi: 10.1016/j.jeconbus.2012.03.002.

[87] L. Tang, C. Zhang, L. Li, and S. Wang, “A multi-scale method for forecasting oil price with multi-factor search engine data,” *Applied Energy*, vol. 257, p. 114033, Jan. 2020, doi: 10.1016/j.apenergy.2019.114033.

[88] Y. Fang, W. Wang, P. Wu, and Y. Zhao, “A sentiment-enhanced hybrid model for crude oil price forecasting,” *Expert Systems with Applications*, vol. 215, p. 119329, Apr. 2023, doi: 10.1016/j.eswa.2022.119329.

[89] W. Medhat, A. Hassan, and H. Korashy, “Sentiment analysis algorithms and applications: A survey,” *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, Dec. 2014, doi: 10.1016/j.asej.2014.04.011.

[90] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif Intell Rev*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022, doi: 10.1007/s10462-022-10144-1.

[91] R. Prabowo and M. Thelwall, "Sentiment analysis: A combined approach," *Journal of Informetrics*, vol. 3, no. 2, pp. 143–157, Apr. 2009, doi: 10.1016/j.joi.2009.01.003.

[92] M. D. Devika, C. Sunitha, and A. Ganesh, "Sentiment Analysis: A Comparative Study on Different Approaches," *Procedia Computer Science*, vol. 87, pp. 44–49, Jan. 2016, doi: 10.1016/j.procs.2016.05.124.

[93] L. I. Tan, W. S. Phang, K. O. Chin, and A. Patricia, "Rule-Based Sentiment Analysis for Financial News," in *2015 IEEE International Conference on Systems, Man, and Cybernetics*, Oct. 2015, pp. 1601–1606. doi: 10.1109/SMC.2015.283.

[94] W. Zhang, Y. Deng, B. Liu, S. J. Pan, and L. Bing, "Sentiment Analysis in the Era of Large Language Models: A Reality Check." arXiv, May 24, 2023. doi: 10.48550/arXiv.2305.15005.

[95] Y. Yang, M. C. S. UY, and A. Huang, "FinBERT: A Pretrained Language Model for Financial Communications." arXiv, Jul. 08, 2020. doi: 10.48550/arXiv.2006.08097.

[96] H. Kaplan, R.-P. Mundani, H. Rölke, and A. Weichselbraun, "CrudeBERT: Applying Economic Theory towards fine-tuning Transformer-based Sentiment Analysis Models to the Crude Oil Market," in *Proceedings of the 25th International Conference on Enterprise Information Systems*, 2023, pp. 324–334. doi: 10.5220/0011749600003467.

[97] A. Vaswani *et al.*, "Attention Is All You Need," Jun. 2017.

[98] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." arXiv, May 24, 2019. doi: 10.48550/arXiv.1810.04805.

[99] A. Radford and K. Narasimhan, "Improving Language Understanding by Generative Pre-Training," 2018. Accessed: Jul. 15, 2023. [Online]. Available: https://www.semanticscholar.org/paper/Improving-Language-Understanding-by-Generative-Radford-Narasimhan/cd18800a0fe0b668a1cc19f2ec95b5003d0a5035

[100] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language Models are Unsupervised Multitask Learners," 2019. Accessed: Jul. 15, 2023. [Online]. Available: https://www.semanticscholar.org/paper/Language-Models-

are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe

[101] T. B. Brown *et al.*, “Language Models are Few-Shot Learners.” arXiv, Jul. 22, 2020. doi: 10.48550/arXiv.2005.14165.

[102] OpenAI, “GPT-4 Technical Report.” arXiv, Mar. 27, 2023. doi: 10.48550/arXiv.2303.08774.

[103] S. Wu *et al.*, “BloombergGPT: A Large Language Model for Finance.” arXiv, May 09, 2023. doi: 10.48550/arXiv.2303.17564.

[104] J. Lee *et al.*, “BioBERT: a pre-trained biomedical language representation model for biomedical text mining,” *Bioinformatics*, vol. 36, no. 4, pp. 1234–1240, Feb. 2020, doi: 10.1093/bioinformatics/btz682.

[105] I. Beltagy, K. Lo, and A. Cohan, “SciBERT: A Pretrained Language Model for Scientific Text.” arXiv, Sep. 10, 2019. doi: 10.48550/arXiv.1903.10676.

[106] H.-T. Nguyen, “A Brief Report on LawGPT 1.0: A Virtual Legal Assistant Based on GPT-3.” arXiv, Feb. 14, 2023. doi: 10.48550/arXiv.2302.05729.

[107] M. Qadan and H. Nama, “Investor sentiment and the price of oil,” *Energy Economics*, vol. 69, pp. 42–58, Jan. 2018, doi: 10.1016/j.eneco.2017.10.035.

[108] P. K. Narayan, “Oil price news and COVID-19—Is there any connection?,” *Energy RESEARCH LETTERS*, vol. 1, no. 1, Jun. 2020, doi: 10.46557/001c.13176.

[109] L. T. He and K. M. Casey, “Forecasting ability of the investor sentiment endurance index: The case of oil service stock returns and crude oil prices,” *Energy Economics*, vol. 47, pp. 121–128, Jan. 2015, doi: 10.1016/j.eneco.2014.11.005.

[110] Z. Hu, “Crude oil price prediction using CEEMDAN and LSTM-attention with news sentiment index,” *Oil & Gas Science and Technology*, vol. 76, p. 28, Jan. 2021, doi: 10.2516/ogst/2021010.

[111] B. Wu, L. Wang, S. Wang, and Y.-R. Zeng, “Forecasting the U.S. oil markets based on social media information during the COVID-19 pandemic,” *Energy*, vol. 226, p. 120403, Jul. 2021, doi: 10.1016/j.energy.2021.120403.

[112] M. Oussalah and A. Zaidi, “Forecasting Weekly Crude Oil Using Twitter Sentiment of U.S. Foreign Policy and Oil Companies Data,” in *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, Jul. 2018, pp. 201–208. doi: 10.1109/IRI.2018.00037.

[113] B. Wu, L. Wang, S.-X. Lv, and Y.-R. Zeng, “Effective crude oil price forecasting using new text-based and big-data-driven model,” *Measurement*, vol. 168, p. 108468, Jan. 2021, doi: 10.1016/j.measurement.2020.108468.

[114] J. Guo, Z. Zhao, J. Sun, and S. Sun, “Multi-perspective crude oil price forecasting with a new decomposition-ensemble framework,” *Resources Policy*, vol. 77, p. 102737, Aug. 2022, doi: 10.1016/j.resourpol.2022.102737.

[115] X. Gong, K. Guan, and Q. Chen, “The role of textual analysis in oil futures price forecasting based on machine learning approach,” *Journal of Futures Markets*, vol. 42, no. 10, pp. 1987–2017, 2022, doi: 10.1002/fut.22367.

[116] H. Kaplan, A. Weichselbraun, and A. Brasoveanu, “Integrating Economic Theory, Domain Knowledge, and Social Knowledge into Hybrid Sentiment Models for Predicting Crude Oil Markets,” *Cognitive Computation*, pp. 1–17, Mar. 2023, doi: 10.1007/s12559-023-10129-4.

[117] C. D. K. II and J. A. Dahlquist, *Technical Analysis: The Complete Resource for Financial Market Technicians*. FT Press, 2010.

[118] S. A. Nelson, *Abc of stock speculation*. Place of publication not identified: Theclassics Us, 2013.

[119] C.-H. Park and S. H. Irwin, “The Profitability of Technical Analysis: A Review.” Rochester, NY, Oct. 01, 2004. doi: 10.2139/ssrn.603481.

[120] L. Bachelier, “Louis Bachelier’s theory of speculation: The origins of modern finance,” *Louis Bachelier’s Theory of Speculation: The Origins of Modern Finance*, pp. 1–188, Jan. 2011.

[121] E. F. Fama, “Random Walks in Stock Market Prices,” *Financial Analysts Journal*, vol. 21, no. 5, pp. 55–59, 1965.

[122] E. F. Fama, “Efficient Capital Markets: A Review of Theory and Empirical Work,” *The Journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970, doi: 10.2307/2325486.

[123] R. T. Farias Nazário, J. L. e Silva, V. A. Sobreiro, and H. Kimura, “A literature review of technical analysis on stock markets,” *The Quarterly Review of Economics and Finance*, vol. 66, pp. 115–126, Nov. 2017, doi: 10.1016/j.qref.2017.01.014.

[124] C.-H. Park and S. H. Irwin, “A reality check on technical trading rule profits in the U.S. futures markets,” *Journal of Futures Markets*, vol. 30, no. 7, pp. 633–659, 2010, doi: 10.1002/fut.20435.

[125] I. Psaradellis, J. Laws, A. A. Pantelous, and G. Sermpinis, “Performance of technical trading rules: evidence from the crude oil market,” *The European Journal of Finance*, vol. 25, no. 17, pp. 1793–1815, Nov. 2019, doi: 10.1080/1351847X.2018.1552172.

[126] R. Sullivan, A. Timmermann, and H. White, “Data-Snooping, Technical Trading Rule Performance, and the Bootstrap,” *The Journal of Finance*, vol. 54, no. 5, pp. 1647–1691, 1999, doi: 10.1111/0022-1082.00163.

[127] C. Tudor and A. Anghel, “The Financialization of Crude Oil Markets and Its Impact on Market Efficiency: Evidence from the Predictive Ability and Performance of Technical Trading Strategies,” *Energies*, vol. 14, p. 4485, Jul. 2021, doi: 10.3390/en14154485.

[128] P. K. Narayan, S. Narayan, and S. S. Sharma, “An analysis of commodity markets: What gain for investors?,” *Journal of Banking & Finance*, vol. 37, no. 10, pp. 3878–3889, Oct. 2013, doi: 10.1016/j.jbankfin.2013.07.009.

[129] X. Jin, “Evaluating the predictive power of intraday technical trading in China’s crude oil market,” *Journal of Forecasting*, vol. 41, no. 7, pp. 1416–1432, 2022, doi: 10.1002/for.2873.

[130] L. Yin and Q. Yang, “Predicting the oil prices: Do technical indicators help?,” *Energy Economics*, vol. 56, pp. 338–350, May 2016, doi: 10.1016/j.eneco.2016.03.017.

[131] L. Liu, Y. Wang, and L. Yang, “Predictability of crude oil prices: An investor perspective,” *Energy Economics*, vol. 75, pp. 193–205, Sep. 2018, doi: 10.1016/j.eneco.2018.08.010.

[132] Y. Zhang, F. Ma, B. Shi, and D. Huang, “Forecasting the prices of crude oil: An iterated combination approach,” *Energy Economics*, vol. 70, pp. 472–483, Feb. 2018, doi: 10.1016/j.eneco.2018.01.027.

[133] Y. Zhang, F. Ma, and Y. Wang, “Forecasting crude oil prices with a large set of predictors: Can LASSO select powerful predictors?,” *Journal of Empirical Finance*, vol. 54, pp. 97–117, Dec. 2019, doi: 10.1016/j.jempfin.2019.08.007.

[134] G. Phillips-Wren, “Intelligent Decision Support Systems,” in *Multicriteria Decision Aid and Artificial Intelligence*, 2013, pp. 25–44. doi: 10.1002/9781118522516.ch2.

[135] D. Kumar, “Data Mining Based Marketing Decision Support System Using Hybrid Machine Learning Algorithm,” *Journal of Artificial Intelligence and Capsule Networks*, vol. 2, pp. 185–193, Aug. 2020, doi: 10.36548//jaicn.2020.3.006.

[136] C. Zopounidis, M. Doumpos, and D. Niklis, “Financial decision support: an overview of developments and recent trends,” *EURO Journal on Decision Processes*, vol. 6, no. 1, pp. 63–76, Jun. 2018, doi: 10.1007/s40070-018-0078-3.

[137] S. Almahdi and S. Y. Yang, “An adaptive portfolio trading system: A risk-return portfolio optimization using recurrent reinforcement learning with expected maximum drawdown,” *Expert Systems with Applications*, vol. 87, pp. 267–279, Nov. 2017, doi: 10.1016/j.eswa.2017.06.023.

[138] S. Patalay and B. M. Rao, *Artificial Intelligence Based System for Financial Decision Support*. 2021.

[139] A. M.Vaidya, N. H. Waghela, and S. S. Yewale, “Decision Support System for the Stock Market using Data Analytics and Artificial Intelligence,” *IJCA*, vol. 117, no. 8, pp. 21–28, May 2015, doi: 10.5120/20574-2977.

[140] C. Y. Chai and S. Mohd Taib, “Hybrid Stock Investment Strategy Decision Support System : Integration of Data Mining, Artificial Intelligence and Decision Support,” S.-I. Ao, B. Rieger, and M. A. Amouzegar, Eds., Netherlands: Springer Netherlands, 2010, pp. 481–493. Accessed: Dec. 29, 2022. [Online]. Available: http://www.springerlink.com/content/978-90-481-9418-6#section=791672&page=1&locus=0

[141] J. Liu and K. Lee, “An Intelligent Business Advisor System for Stock Investment,” *Expert Systems*, vol. 14, pp. 129–139, Aug. 1997, doi: 10.1111/1468-0394.00049.

[142] R. J. Kuo, C. H. Chen, and Y. C. Hwang, “An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network,” *Fuzzy Sets and Systems*, vol. 118, no. 1, pp. 21–45, 2001, doi: 10.1016/S0165-0114(98)00399-6.

[143] T. Anbalagan and S. U. Maheswari, “Classification and Prediction of Stock Market Index Based on Fuzzy Metagraph,” *Procedia Computer Science*, vol. 47, pp. 214–221, Dec. 2015, doi: 10.1016/j.procs.2015.03.200.

[144] V. Cho, “MISMIS – A comprehensive decision support system for stock market investment,” *Knowl Based Syst*, vol. 23, no. 6, pp. 626–633, Aug. 2010, doi: 10.1016/j.knosys.2010.04.009.

[145] Y. Sun, M. Fang, and X. Wang, “A novel stock recommendation system using Guba sentiment analysis,” *Pers Ubiquit Comput*, vol. 22, no. 3, pp. 575–587, Jun. 2018, doi: 10.1007/s00779-018-1121-x.

[146] J. G. Davis, E. Subrahmanian, and A. W. Westerberg, "SCOPE: a blackboard model-based decision support system for crude-oil trading," *Intelligent Systems in Accounting, Finance and Management*, vol. 8, no. 2, pp. 89–104, 1999.

[147] C. Tudor and R. Sova, "Flexible Decision Support System for Algorithmic Trading: Empirical Application on Crude Oil Markets," *IEEE Access*, vol. 10, pp. 9628–9644, 2022, doi: 10.1109/ACCESS.2022.3143767.

[148] C.-F. Cheng and C. J. Lin, "Building a Low-Cost Wireless Biofeedback Solution: Applying Design Science Research Methodology," *Sensors*, vol. 23, no. 6, 2023, doi: 10.3390/s23062920.

[149] K. Peffers, T. Tuunanen, M. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *Journal of Management Information Systems*, vol. 24, pp. 45–77, Jan. 2007.

[150] A. M., M. Zekiwos, and A. Bruck, "Deep Learning-Based Image Processing for Cotton Leaf Disease and Pest Diagnosis," *Journal of Electrical and Computer Engineering*, vol. 2021, p. e9981437, Jun. 2021, doi: 10.1155/2021/9981437.

[151] D. Kudryavtsev and T. Gavrilova, "From Anarchy to System: A Novel Classification of Visual Knowledge Codification Techniques," *Knowledge and Process Management*, vol. 24, no. 1, 2016, doi: 10.1002/kpm.1509.

[152] J. R. Venable, J. Pries-Heje, and R. L. Baskerville, "FEDS: a Framework for Evaluation in Design Science Research," *European Journal of Information Systems*, vol. 25, no. 1, 2016, doi: 10.1057/ejis.2014.36.

[153] A. Hevner and S. Chatterjee, "Design Science Research in Information Systems," in *Management Information Systems Quarterly - MISQ*, 2010, pp. 9–22. doi: 10.1007/978-1-4419-5653-8_2.

[154] B. Kitchenham, "Procedures for Performing Systematic Reviews," *Keele, UK, Keele Univ.*, vol. 33, Aug. 2004.

[155] Y. Zhao, J. Li, and L. Yu, "A deep learning ensemble approach for crude oil price forecasting," *Energy Economics*, vol. 66, pp. 9–16, Aug. 2017, doi: 10.1016/j.eneco.2017.05.023.

[156] X. Ding, L. Fu, Y. Ding, and Y. Wang, "A novel hybrid method for oil price forecasting with ensemble thought," *Energy Reports*, vol. 8, pp. 15365–15376, Nov. 2022, doi: 10.1016/j.egyr.2022.11.061.

[157] L. Gabralla and A. Abraham, "Comparison of Hybrid Intelligent Approaches for Prediction of Crude Oil Price," *International Journal of Computer Information Systems and Industrial Management Applications 2150-7988*, vol. 7, Dec. 2015.

[158] M. Fouladgar, M. Yazdani, S. Khazaee, E. Zavadskas, and V. Fouladgar, "Comparison of vector time series and ann techniques for forecasting of wti oil price," *Economic computation and economic cybernetics studies and research / Academy of Economic Studies*, vol. 47, pp. 19–34, Jan. 2013.

[159] K. C. S, S. Mahadevan, and S. N. Sivanandam, "Crude oil prediction using a hybrid radial basis function network," *Journal of Theoretical and Applied Information Technology*, vol. 72, pp. 199–207, Jan. 2015.

[160] Y. Bai, X. Li, H. Yu, and S. Jia, "Crude oil price forecasting incorporating news text," *International Journal of Forecasting*, vol. 38, no. 1, pp. 367–383, Jan. 2022, doi: 10.1016/j.ijforecast.2021.06.006.

[161] Y. Pang *et al.*, "Forecasting the crude oil spot price by wavelet neural networks using oecd petroleum inventory levels," *New Math. and Nat. Computation*, vol. 07, no. 02, pp. 281–297, May 2011, doi: 10.1142/S1793005711001937.

[162] Y. Yang, J. Guo, S. Sun, and Y. Li, "Forecasting crude oil price with a new hybrid approach and multi-source data," *Engineering Applications of Artificial Intelligence*, vol. 101, p. 104217, May 2021, doi: 10.1016/j.engappai.2021.104217.

[163] J. Li, L. Tang, and S. Wang, "Forecasting crude oil price with multilingual search engine data," *Physica A: Statistical Mechanics and its Applications*, vol. 551, p. 124178, Aug. 2020, doi: 10.1016/j.physa.2020.124178.

[164] S. Ramyar and F. Kianfar, "Forecasting Crude Oil Prices: A Comparison Between Artificial Neural Networks and Vector Autoregressive Models," *Comput Econ*, vol. 53, no. 2, pp. 743–761, Feb. 2019, doi: 10.1007/s10614-017-9764-7.

[165] L.-T. Zhao, G.-R. Zeng, W.-J. Wang, and Z.-G. Zhang, "Forecasting Oil Price Using Web-based Sentiment Analysis," *Energies*, vol. 12, p. 4291, Nov. 2019, doi: 10.3390/en12224291.

[166] X. Hao, Y. Zhao, and Y. Wang, "Forecasting the real prices of crude oil using robust regression models with regularization constraints," *Energy Economics*, vol. 86, p. 104683, Feb. 2020, doi: 10.1016/j.eneco.2020.104683.

[167] L.-T. Zhao, Z.-Y. Zheng, Y. Fu, Z.-X. Liu, and M.-F. Li, "Google Index-Driven Oil Price Value-at-Risk Forecasting: A Decomposition Ensemble Approach," *IEEE Access*, vol. 8, pp. 183351–183366, 2020, doi: 10.1109/ACCESS.2020.3028124.

[168] M. Abd Elaziz, A. A. Ewees, and Z. Alameer, "Improving Adaptive Neuro-Fuzzy Inference System Based on a Modified Salp Swarm Algorithm Using Genetic

Algorithm to Forecast Crude Oil Price," *Nat Resour Res*, vol. 29, no. 4, pp. 2671–2686, Aug. 2020, doi: 10.1007/s11053-019-09587-1.

[169] J. Li, S. Zhu, and Q. Wu, "Monthly crude oil spot price forecasting using variational mode decomposition," *Energy Economics*, vol. 83, pp. 240–253, Sep. 2019, doi: 10.1016/j.eneco.2019.07.009.

[170] H. Chiroma and S. Abdul Kareem, "420-Neuro–Genetic Model for Crude Oil Price Prediction while Considering the Impact of Uncertainties," *International Journal of Oil Gas and Coal Technology*, vol. 12, Jan. 2015, doi: 10.1504/IJOGCT.2016.076803.

[171] E. Lotfi and M. Karimi, "Opec Oil Price Prediction Using Anfis," *Journal of Mathematics and Computer Science*, vol. 10, pp. 286–296, May 2014, doi: 10.22436/jmcs.010.04.07.

[172] K. Bekiroglu, O. Duru, E. Gulay, R. Su, and C. Lagoa, "Predictive analytics of crude oil prices by utilizing the intelligent model search engine," *Applied Energy*, vol. 228, pp. 2387–2397, Oct. 2018, doi: 10.1016/j.apenergy.2018.07.071.

[173] X. Li, W. Shang, and S. Wang, "Text-based crude oil price forecasting: A deep learning approach," *International Journal of Forecasting*, vol. 35, no. 4, pp. 1548–1560, Oct. 2019, doi: 10.1016/j.ijforecast.2018.07.006.

[174] K. A. Aastveit, H. C. Bjï¿½rnland, and L. Thorsrud, "What drives oil prices? Emerging versus developed economies," Centre for Applied Macro- and Petroleum economics (CAMP), BI Norwegian Business School, Working Paper No 2/2012, Dec. 2012. Accessed: Jul. 03, 2023. [Online]. Available: https://econpapers.repec.org/paper/bnywpaper/0007.htm

[175] A. Adedeji, F. Ahmed, and M. Muhammed, "Empirical Evidence Of Oil Price Shocks And Oil Economy Asymmetric Nexus: The Cases Of Angola And Nigeria," *Advances in Social Sciences Research Journal*, Jun. 2018, doi: 10.14738/assrj.56.4649.

[176] J. Hamilton and A. M. Herrera, "Oil Shocks and Aggregate Macroeconomic Behavior: The Role of Monetary Policy: Comment," *Journal of Money, Credit and Banking*, vol. 36, no. 2, pp. 265–86, 2004.

[177] J. Matres and T. Le, "The Impact of Money Supply on the Economy: A Panel Study on Selected Countries," *Journal of Economic Science Research*, vol. 4, Oct. 2021, doi: 10.30564/jesr.v4i4.3782.

[178] G. Kumar, U. P. Singh, and S. Jain, "Hybrid evolutionary intelligent system and hybrid time series econometric model for stock price forecasting," *International*

*Journal of Intelligent Systems*, vol. 36, no. 9, pp. 4902–4935, 2021, doi: 10.1002/int.22495.

[179] R. Arévalo, J. García, F. Guijarro, and A. Peris, "A dynamic trading rule based on filtered flag pattern recognition for stock market price forecasting," *Expert Systems with Applications*, vol. 81, pp. 177–192, Sep. 2017, doi: 10.1016/j.eswa.2017.03.028.

[180] E. Elsayed and F. Ghanam, "SGEDSS: Semantic Gene Expression Model for Communication Decision Supportsystem," *IOSR Journal of Computer Engineering*, vol. 18, pp. 46–52, Apr. 2016, doi: 10.9790/0661-1804044652.

[181] N. Khan *et al.*, "Big Data: Survey, Technologies, Opportunities, and Challenges," *The Scientific World Journal*, vol. 2014, p. e712826, Jul. 2014, doi: 10.1155/2014/712826.

[182] N. Elgendy, A. Elragal, and T. Päivärinta, "DECAS: a modern data-driven decision theory for big data and analytics," *Journal of Decision Systems*, vol. 31, no. 4, pp. 337–373, Oct. 2022, doi: 10.1080/12460125.2021.1894674.

[183] Y. Demchenko, P. Grosso, C. de Laat, and P. Membrey, "Addressing big data issues in Scientific Data Infrastructure," in *2013 International Conference on Collaboration Technologies and Systems (CTS)*, May 2013, pp. 48–55. doi: 10.1109/CTS.2013.6567203.

[184] K. Jee and G.-H. Kim, "Potentiality of Big Data in the Medical Sector: Focus on How to Reshape the Healthcare System," *Healthc Inform Res*, vol. 19, no. 2, pp. 79–85, Jun. 2013, doi: 10.4258/hir.2013.19.2.79.

[185] C. Yang, Q. Huang, Z. Li, K. Liu, and F. Hu, "Big Data and cloud computing: innovation opportunities and challenges," *International Journal of Digital Earth*, vol. 10, no. 1, pp. 13–53, Jan. 2017, doi: 10.1080/17538947.2016.1239771.

[186] M. S. Mahmud, J. Z. Huang, S. Salloum, T. Z. Emara, and K. Sadatdiynov, "A survey of data partitioning and sampling methods to support big data analysis," *Big Data Mining and Analytics*, vol. 3, no. 2, pp. 85–101, Jun. 2020, doi: 10.26599/BDMA.2019.9020015.

[187] J. Zhang and M. Lin, "A comprehensive bibliometric analysis of Apache Hadoop from 2008 to 2020," *International Journal of Intelligent Computing and Cybernetics*, vol. 16, no. 1, pp. 99–120, Jan. 2022, doi: 10.1108/IJICC-01-2022-0004.

[188] P. R. Giri and G. Sharma, "Apache Hadoop Architecture, Applications, and Hadoop Distributed File System," *Semiconductor Science and Information Devices*, vol. 4, no. 1, Art. no. 1, May 2022, doi: 10.30564/ssid.v4i1.4619.

[189] A. Bayar *et al.*, “A Comprehensive Big Data Framework for Energy Markets: Oil and Gas Demand Forecasting,” in *2022 3rd International Informatics and Software Engineering Conference (IISEC)*, Dec. 2022, pp. 1–6. doi: 10.1109/IISEC56263.2022.9998216.

[190] N. V. Patil, C. R. Krishna, and K. Kumar, “KS-DDoS: Kafka streams-based classification approach for DDoS attacks,” *J Supercomput*, vol. 78, no. 6, pp. 8946–8976, Apr. 2022, doi: 10.1007/s11227-021-04241-1.

[191] H. Zhang, “Research and Design of Real Time Big Data Visualization Analysis Platform Based on Flink,” *J. Phys.: Conf. Ser.*, vol. 2504, no. 1, p. 012053, May 2023, doi: 10.1088/1742-6596/2504/1/012053.

[192] M. Syafrudin, G. Alfian, N. L. Fitriyani, and J. Rhee, “Performance Analysis of IoT-Based Sensor, Big Data Processing, and Machine Learning Model for Real-Time Monitoring System in Automotive Manufacturing,” *Sensors*, vol. 18, no. 9, Art. no. 9, Sep. 2018, doi: 10.3390/s18092946.

[193] T. R. Aravindan, C. N. Vigneshwar, and S. G, “Sentiment Classification for Amazon Fine Foods Reviews Using Pyspark,” in *Recent Developments in Electronics and Communication Systems*, IOS Press, 2023, pp. 431–436. doi: 10.3233/ATDE221293.

[194] K. Vyas, P. M. Jat, K. Vyas, and P. M. Jat, “Study of Consistency and Performance Trade-Off in Cassandra,” *Computer Science & Information Technology (CS & IT)*, vol. 12, no. 19, Art. no. 19, Nov. 2022, doi: 10.5121/csit.2022.121907.

[195] M. V. Santos, F. Morgado-Dias, and T. C. Silva, “Oil Sector and Sentiment Analysis—A Review,” *Energies*, vol. 16, no. 12, Art. no. 12, Jan. 2023, doi: 10.3390/en16124824.

[196] S. Ben Jabeur, R. Khalfaoui, and W. Ben Arfi, “The effect of green energy, global environmental indexes, and stock markets in predicting oil price crashes: Evidence from explainable machine learning,” *J Environ Manage*, vol. 298, p. 113511, Nov. 2021, doi: 10.1016/j.jenvman.2021.113511.

[197] A. P. Engelbrecht, *Computational Intelligence*. Chichester, UK: John Wiley & Sons, Ltd, 2007. doi: 10.1002/9780470512517.

[198] Y. Zhao and J. Xiao, “An adiabatic method to train binarized artificial neural networks,” *Sci Rep*, vol. 11, no. 1, Art. no. 1, Oct. 2021, doi: 10.1038/s41598-021-99191-2.

[199] J. Siswantoro, A. S. Prabuwono, A. Abdullah, and B. Idrus, “A linear model based on Kalman filter for improving neural network classification performance,”

*Expert Systems with Applications*, vol. 49, pp. 112–122, May 2016, doi: 10.1016/j.eswa.2015.12.012.

[200] H. Hewamalage, C. Bergmeir, and K. Bandara, “Recurrent Neural Networks for Time Series Forecasting: Current status and future directions,” *International Journal of Forecasting*, vol. 37, no. 1, pp. 388–427, Jan. 2021, doi: 10.1016/j.ijforecast.2020.06.008.

[201] G. Petneházi, *Recurrent Neural Networks for Time Series Forecasting*. 2018.

[202] S. Hochreiter and J. Schmidhuber, “Long Short-term Memory,” *Neural computation*, vol. 9, pp. 1735–80, Dec. 1997, doi: 10.1162/neco.1997.9.8.1735.

[203] H. Jiang, P. He, W. Chen, X. Liu, J. Gao, and T. Zhao, *SMART: Robust and Efficient Fine-Tuning for Pre-trained Natural Language Models through Principled Regularized Optimization*. 2020, p. 2190. doi: 10.18653/v1/2020.acl-main.197.

[204] B. Zhang, H. Yang, and X.-Y. Liu, “Instruct-FinGPT: Financial Sentiment Analysis by Instruction Tuning of General-Purpose Large Language Models.” arXiv, Jun. 21, 2023. doi: 10.48550/arXiv.2306.12659.

[205] A. Lopez-Lira and Y. Tang, “Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models.” Rochester, NY, Apr. 06, 2023. doi: 10.2139/ssrn.4412788.

[206] M. Ali, M. Jones, X. Xie, and M. Williams, “Towards Visual Exploration of Large Temporal Datasets,” in *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, Oct. 2018, pp. 1–9. doi: 10.1109/BDVA.2018.8534025.

[207] M. Taylor and H. Allen, “The use of technical analysis in the foreign exchange market,” *Journal of International Money and Finance*, vol. 11, no. 3, pp. 304–314, 1992.

[208] P.-H. Hsu, M. P. Taylor, and Z. Wang, “Technical trading: Is it still beating the foreign exchange market?,” *Journal of International Economics*, vol. 102, pp. 188–208, Sep. 2016, doi: 10.1016/j.jinteco.2016.03.012.

[209] R. Hudson and A. Urquhart, “Technical trading and cryptocurrencies,” *Ann Oper Res*, vol. 297, no. 1, pp. 191–220, Feb. 2021, doi: 10.1007/s10479-019-03357-1.

[210] J. Oh, “Short-Term Stock Price Prediction by Supervised Learning of Rapid Volume Decrease Patterns,” *IEICE Transactions on Information and Systems*, vol. E105.D, no. 8, pp. 1431–1442, 2022, doi: 10.1587/transinf.2021EDP7243.

[211] G. Liu, “Technical Trading Behaviour: Evidence from Chinese Rebar Futures Market,” *Comput Econ*, vol. 54, no. 2, pp. 669–704, Aug. 2019, doi: 10.1007/s10614-018-9851-4.

[212] I. M. A. I. Agusta, A. Barakbah, and A. Fariza, “Technical Analysis Based Automatic Trading Prediction System for Stock Exchange using Support Vector Machine,” *EMITTER International Journal of Engineering Technology*, pp. 279–293, Dec. 2022, doi: 10.24003/emitter.v10i2.740.

[213] T. N. Bulkowski, *Encyclopedia of candlestick charts*. Hoboken, N.J.: J. Wiley & Sons, 2008. Accessed: Aug. 04, 2023. [Online]. Available: http://site.ebrary.com/id/10226911

[214] J.-H. Chen and Y.-C. Tsai, “Encoding candlesticks as images for pattern classification using convolutional neural networks,” *Financial Innovation*, vol. 6, no. 1, p. 26, Jun. 2020, doi: 10.1186/s40854-020-00187-0.

[215] P. L. Adhikari, “The dynamic relationship between stock returns and trading volumes in Nepalese stock market,” *Management Dynamics*, vol. 23, no. 2, Art. no. 2, Dec. 2020, doi: 10.3126/md.v23i2.35819.

[216] W. Lertthaweedech, P. Kantavat, and B. Kijsirikul, “Effective Crude Oil Trading Techniques Using Long Short-Term Memory and Convolution Neural Networks,” *Journal of Advances in Information Technology*, vol. 13, no. 6, pp. 645–651, 2022, doi: 10.12720/jait.13.6.645-651.

[217] M. Zatwarnicki, K. Zatwarnicki, and P. Stolarski, “Effectiveness of the Relative Strength Index Signals in Timing the Cryptocurrency Market,” *Sensors*, vol. 23, no. 3, Art. no. 3, Jan. 2023, doi: 10.3390/s23031664.

[218] H. Tadas, J. Nagarkar, S. Malik, D. K. Mishra, and D. Paul, “The effectiveness of technical trading strategies: Evidence from Indian equity markets,” *Investment Management and Financial Innovations*, vol. 20, no. 2, pp. 26–40, Apr. 2023, doi: 10.21511/imfi.20(2).2023.03.

[219] D. L. Joshi, “Use of Moving Average Convergence Divergence for Predicting Price Movements,” *International Research Journal of MMC*, vol. 3, no. 4, Art. no. 4, Oct. 2022, doi: 10.3126/irjmmc.v3i4.48859.

[220] I. Gurrib, M. Nourani, and R. K. Bhaskaran, “Energy crypto currencies and leading U.S. energy stock prices: are Fibonacci retracements profitable?,” *Financial Innovation*, vol. 8, no. 1, p. 8, Jan. 2022, doi: 10.1186/s40854-021-00311-8.

[221] J. Chakole and M. Kurhekar, “Trend following deep Q-Learning strategy for stock trading,” *Expert Systems*, vol. 37, no. 4, p. e12514, 2020, doi: 10.1111/exsy.12514.

[222] G. Cohen, “Trading cryptocurrencies using algorithmic average true range systems,” *Journal of Forecasting*, vol. 42, no. 2, pp. 212–222, 2023, doi: 10.1002/for.2906.

[223] “Dash Documentation & User Guide | Plotly.” https://dash.plotly.com/ (accessed Aug. 10, 2023).

[224] Y. Li, S. Jiang, X. Li, and S. Wang, “The role of news sentiment in oil futures returns and volatility forecasting: Data-decomposition based deep learning approach,” *Energy Economics*, vol. 95, no. C, 2021, Accessed: Aug. 06, 2023. [Online]. Available: https://ideas.repec.org//a/eee/eneeco/v95y2021ics0140988321000451.html

[225] Y. Xu and R. Xu, “Research on Interpolation and Data Fitting: Basis and Applications.” arXiv, Aug. 24, 2022. doi: 10.48550/arXiv.2208.11825.

[226] H. Henderi, T. Wahyuningsih, and E. Rahwanto, “Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer,” *International Journal of Informatics and Information Systems*, vol. 4, no. 1, Art. no. 1, Mar. 2021, doi: 10.47738/ijiis.v4i1.73.

[227] F. T. Liu, K. M. Ting, and Z.-H. Zhou, “Isolation Forest,” in *2008 Eighth IEEE International Conference on Data Mining*, Dec. 2008, pp. 413–422. doi: 10.1109/ICDM.2008.17.

[228] N. E. Huang *et al.*, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, Mar. 1998, doi: 10.1098/rspa.1998.0193.

[229] J. L. Bettman, S. J. Sault, and E. L. Schultz, “Fundamental and technical analysis: substitutes or complements?,” *Accounting & Finance*, vol. 49, no. 1, pp. 21–36, 2009, doi: 10.1111/j.1467-629X.2008.00277.x.

[230] S. Gupta, “FUNDAMENTAL & TECHNICAL ANAYSIS OF CRUDE OIL PRICES,” *Journal of Global Economy*, vol. 17, no. 1, Art. no. 1, Apr. 2021, doi: 10.1956/jge.v17i1.617.

[231] I. K. Nti, A. F. Adekoya, and B. A. Weyori, “A systematic review of fundamental and technical analysis of stock market predictions,” *Artif Intell Rev*, vol. 53, no. 4, pp. 3007–3057, Apr. 2020, doi: 10.1007/s10462-019-09754-z.

[232] W.-C. Chiang, D. Enke, T. Wu, and R. Wang, “An adaptive stock index trading decision support system,” *Expert Systems with Applications*, vol. 59, pp. 195–207, Oct. 2016, doi: 10.1016/j.eswa.2016.04.025.

[233] C. Tudor and R. Sova, “Flexible Decision Support System for Algorithmic Trading: Empirical Application on Crude Oil Markets,” *IEEE Access*, vol. 10, pp. 9628–9644, 2022, doi: 10.1109/ACCESS.2022.3143767.

# APPENDICES

## APPENDIX A

## DATA FOR CRUDE OIL PRICE FORECASTING

Table 7: Data for Crude Oil Price Forecasting from SLR Results

| Category | Name |
|---|---|
| Consumption, orders, and inventory | Consumer Sentiment Index |
| Consumption, orders, and inventory | New Orders for Consumer Goods |
| Consumption, orders, and inventory | New Orders for Durable Goods |
| Consumption, orders, and inventory | New Orders for Nondefense Capital Goods |
| Consumption, orders, and inventory | Real Manu. and Trade Industries Sales |
| Consumption, orders, and inventory | Real personal consumption expenditures |
| Consumption, orders, and inventory | Personal Cons. Exp: Durable goods |
| Consumption, orders, and inventory | Personal Cons. Exp: Nondurable goods |
| Consumption, orders, and inventory | Personal Cons. Exp: Services |
| Consumption, orders, and inventory | Personal Cons. Expend.: Chain Index |
| Consumption, orders, and inventory | Retail and Food Services Sales |
| Consumption, orders, and inventory | Total Business Inventories |
| Consumption, orders, and inventory | Total Business: Inventories to Sales Ratio |
| Consumption, orders, and inventory | Unfilled Orders for Durable Goods |
| Economic uncertainty | Canada Policy related economic uncertainty index |
| Economic uncertainty | China Policy related economic uncertainty index |
| Economic uncertainty | Equity Market-Related Economic Uncertainty Index |
| Economic uncertainty | Global Policy related economic uncertainty index (GDP) |
| Economic uncertainty | Global Policy related economic uncertainty index (PPP) |
| Economic uncertainty | Japan Policy related economic uncertainty index |
| Economic uncertainty | Russia Policy related economic uncertainty index |
| Economic uncertainty | UK Policy related economic uncertainty index |
| Economic uncertainty | US Policy related economic uncertainty index |
| Energy Outlook | Baltic Exchange Dirty Tanker Index |
| Energy Outlook | Baltic Exchange Dry Index (BDI) |
| Energy Outlook | Active Well Service Rig Count |
| Energy Outlook | Aircraft Utilization, U.S. (revenue ton miles/day thousands) |
| Energy Outlook | All Grades of Gasoline, U.S. City Average Retail Price |
| Energy Outlook | Asphalt and Road Oil Product Supplied |
| Energy Outlook | Average Free on Board Cost of Crude Oil Imports From All Non-OPEC Countries |
| Energy Outlook | Average Free on Board Cost of Crude Oil Imports From All OPEC Countries |
| Energy Outlook | Average Free on Board Cost of Crude Oil Imports From Colombia |
| Energy Outlook | Average Free on Board Cost of Crude Oil Imports From Mexico |
| Energy Outlook | Average Free on Board Cost of Crude Oil Imports From Nigeria |
| Energy Outlook | Average Free on Board Cost of Crude Oil Imports From Persian Gulf Nations |
| Energy Outlook | Average Free on Board Cost of Crude Oil Imports From Saudi Arabia |
| Energy Outlook | Average Free on Board Cost of Crude Oil Imports From United Kingdom |
| Energy Outlook | Average Free on Board Cost of Crude Oil Imports From Venezuela |
| Energy Outlook | Average Free on Board Costof Crude Oil Imports From Angola |
| Energy Outlook | Average Landed Cost of Crude Oil Imports From All Non-OPEC Countries |
| Energy Outlook | Average Landed Cost of Crude Oil Imports From All OPEC Countries |
| Energy Outlook | Average Landed Cost of Crude Oil Imports From Angola |
| Energy Outlook | Average Landed Cost of Crude Oil Imports From Canada |
| Energy Outlook | Average Landed Cost of Crude Oil Imports From Colombia |
| Energy Outlook | Average Landed Cost of Crude Oil Imports From Mexico |
| Energy Outlook | Average Landed Cost of Crude Oil Imports From Nigeria |
| Energy Outlook | Average Landed Cost of Crude Oil Imports From Persian Gulf Nations |
| Energy Outlook | Average Landed Cost of Crude Oil Imports From Saudi Arabia |
| Energy Outlook | Average Landed Cost of Crude Oil Imports From United Kingdom |
| Energy Outlook | Average Landed Cost of Crude Oil Imports From Venezuela |
| Energy Outlook | Average Refiner Price of Finished Aviation Gasoline for Resale |
| Energy Outlook | Average Refiner Price of Finished Aviation Gasoline to End Users |

| | |
|---|---|
| Energy Outlook | Average Refiner Price of Residual Fuel Oil, Sulfur Content Less Than or Equal to 1 Percent, Sales to End Users |
| Energy Outlook | Aviation Gasoline Product Supplied |
| Energy Outlook | Baker Hughes International Rig Count |
| Energy Outlook | Biofuels Consumption |
| Energy Outlook | Biofuels Production |
| Energy Outlook | Central Atlantic (PADD 1B) Ending Stocks of Distillate Fuel Oil |
| Energy Outlook | Central Atlantic (PADD 1B) Ending Stocks of Total Gasoline |
| Energy Outlook | Changes in Oil Inventory |
| Energy Outlook | China Crude Oil Import |
| Energy Outlook | Coal Consumption, U.S. (million short tons) |
| Energy Outlook | Coal Production, U.S. (million short tons) |
| Energy Outlook | Consumption of Electricity, U.S. (billion kilowatt-hours) |
| Energy Outlook | Cost of Coal Delivered to Electric Generating Plants, U.S. (dollars per million Btu) |
| Energy Outlook | Crude Oil and Natural Gas Rotary Rigs in Operation, Offshore |
| Energy Outlook | Crude Oil and Natural Gas Rotary Rigs in Operation, Onshore |
| Energy Outlook | Crude Oil and Natural Gas Rotary Rigs in Operation, Total |
| Energy Outlook | Crude Oil Exports |
| Energy Outlook | Crude Oil Imports, Total |
| Energy Outlook | Crude Oil Production, Algeria |
| Energy Outlook | Crude Oil Production, Angola |
| Energy Outlook | Crude Oil Production, Canada |
| Energy Outlook | Crude Oil Production, China |
| Energy Outlook | Crude Oil Production, Ecuador |
| Energy Outlook | Crude Oil Production, Egypt |
| Energy Outlook | Crude Oil Production, Indonesia |
| Energy Outlook | Crude Oil Production, Iran |
| Energy Outlook | Crude Oil Production, Kuwait |
| Energy Outlook | Crude Oil Production, Libya |
| Energy Outlook | Crude Oil Production, Mexico |
| Energy Outlook | Crude Oil Production, Nigeria |
| Energy Outlook | Crude Oil Production, Norway |
| Energy Outlook | Crude Oil Production, Persian Gulf Nations |
| Energy Outlook | Crude Oil Production, Qatar |
| Energy Outlook | Crude Oil Production, Russia |
| Energy Outlook | Crude Oil Production, Saudi Arabia |
| Energy Outlook | Crude Oil Production, Total Non-OPEC |
| Energy Outlook | Crude Oil Production, Total OPEC |
| Energy Outlook | Crude Oil Production, United Arab Emirates |
| Energy Outlook | Crude Oil Production, United Kingdom |
| Energy Outlook | Crude Oil Production, United States |
| Energy Outlook | Crude Oil Production, World |
| Energy Outlook | Crude Oil Refinery and Blender Net Input |
| Energy Outlook | Crude Oil Rotary Rigs in Operation |
| Energy Outlook | Crude Oil Stocks at Refineries |
| Energy Outlook | Crude Oil Stocks, Non-Strategic Petroleum Reserve, End of Period (Non SPR) |
| Energy Outlook | Crude Oil Stocks, Strategic Petroleum Reserve, End of Period (SPR) |
| Energy Outlook | Crude Oil Stocks, Total, End of Period |
| Energy Outlook | Crude OilProduction, Iraq |
| Energy Outlook | Crude OilProduction, Venezuela |
| Energy Outlook | Crude Oil Production Capacity, OPEC (million barrels per day) |
| Energy Outlook | Distillate Fuel Oil Imports |
| Energy Outlook | Distillate Fuel Oil Product Supplied |
| Energy Outlook | Distillate Fuel Oil Refinery and Blender Net Production |
| Energy Outlook | Distillate Fuel Oil Stocks, End of Period |
| Energy Outlook | Electricity Retail Sales to the Commercial Sector |
| Energy Outlook | Electricity Retail Sales to the Industrial Sector |
| Energy Outlook | Electricity Retail Sales to the Residential Sector |
| Energy Outlook | Electricity Retail Sales to the Transportation Sector |
| Energy Outlook | Finished Motor Gasoline Imports |
| Energy Outlook | Finished Motor Gasoline Refinery and Blender Net Production |
| Energy Outlook | Freight Rate |
| Energy Outlook | Fuel Ethanol Consumption |
| Energy Outlook | Fuel Ethanol Stocks, End of Period |
| Energy Outlook | Geothermal Energy Consumption/Production |
| Energy Outlook | Global Crude Oil Demand |
| Energy Outlook | Global Crude Oil Stocks |
| Energy Outlook | Jet Fuel Consumed by the Transportation Sector |
| Energy Outlook | Jet Fuel Imports |
| Energy Outlook | Jet Fuel Product Supplied |
| Energy Outlook | Jet Fuel Refinery and Blender Net Production |

| | |
|---|---|
| Energy Outlook | Jet Fuel Stocks, End of Period |
| Energy Outlook | Liquefied Petroleum Gases Imports |
| Energy Outlook | Liquefied Petroleum Gases Product Supplied |
| Energy Outlook | Liquefied Petroleum Gases Refinery and Blender Net Production |
| Energy Outlook | Liquefied Petroleum Gases Stocks, End of Period |
| Energy Outlook | Liquid Fuels Consumption, non OECD (million barrels per day) |
| Energy Outlook | Liquid Fuels Consumption, OECD (million barrels per day) |
| Energy Outlook | Liquid Fuels Consumption, World (million barrels per day) |
| Energy Outlook | Lower Atlantic (PADD 1C) Ending Stocks of Total Gasoline |
| Energy Outlook | Lubricants Product Supplied |
| Energy Outlook | Motor Gasoline Product Supplied |
| Energy Outlook | Motor Gasoline Stocks (Including Blending Components and Gasohol), End of Period |
| Energy Outlook | Natural Gas Plant Liquids Refinery and Blender Net Inputs |
| Energy Outlook | Natural Gas Rotary Rigs in Operation |
| Energy Outlook | Net Inventory Withdrawals, Crude Oil and Other Liquids, U.S. (million barrels per day) |
| Energy Outlook | New England (PADD 1A) Ending Stocks of Distillate Fuel Oil |
| Energy Outlook | New England (PADD 1A) Ending Stocks of Total Gasoline |
| Energy Outlook | Nuclear Electric Power Consumed by the Electric Power Sector |
| Energy Outlook | OECD Total Oil Gross/Net Imports |
| Energy Outlook | OPEC surplus crude oil production capacity |
| Energy Outlook | Other Liquids Refinery and Blender Net Inputs |
| Energy Outlook | Other Petroleum Products Consumed by the Industrial Sector |
| Energy Outlook | Other Petroleum Products Imports |
| Energy Outlook | Other Petroleum Products Refinery and Blender Net Production |
| Energy Outlook | Other Petroleum Products Stocks, End of Period |
| Energy Outlook | Petroleum Coke Product Supplied |
| Energy Outlook | Petroleum Consumption, Canada |
| Energy Outlook | Petroleum Consumption, France |
| Energy Outlook | Petroleum Consumption, Germany |
| Energy Outlook | Petroleum Consumption, Italy |
| Energy Outlook | Petroleum Consumption, Japan |
| Energy Outlook | Petroleum Consumption, OECD Europe |
| Energy Outlook | Petroleum Consumption, Other OECD |
| Energy Outlook | Petroleum Consumption, South Korea |
| Energy Outlook | Petroleum Consumption, Total OECD |
| Energy Outlook | Petroleum Consumption, United Kingdom |
| Energy Outlook | Petroleum Consumption, World |
| Energy Outlook | Petroleum Products Exports |
| Energy Outlook | Petroleum Products Supplied/Consumption, United States |
| Energy Outlook | Petroleum Stocks, Canada, End of Period |
| Energy Outlook | Petroleum Stocks, France, End of Period |
| Energy Outlook | Petroleum Stocks, Germany, End of Period |
| Energy Outlook | Petroleum Stocks, Italy, End of Period |
| Energy Outlook | Petroleum Stocks, Japan, End of Period |
| Energy Outlook | Petroleum Stocks, OECD Europe, End of Period |
| Energy Outlook | Petroleum Stocks, Other OECD, End of Period |
| Energy Outlook | Petroleum Stocks, South Korea, End of Period |
| Energy Outlook | Petroleum Stocks, Total OECD, End of Period |
| Energy Outlook | Petroleum Stocks, United Kingdom, End of Period |
| Energy Outlook | Petroleum Stocks, United States, End of Period |
| Energy Outlook | Propane/Propylene Imports |
| Energy Outlook | Propane/Propylene Product Supplied |
| Energy Outlook | Propane/Propylene Refinery and Blender Net Production |
| Energy Outlook | Propane/Propylene Stocks, End of Period |
| Energy Outlook | Raw Steel Production, U.S. (million short tons per day) |
| Energy Outlook | Refinery Utilization Rate |
| Energy Outlook | Residential Sector Electrical System EnergyLosses |
| Energy Outlook | Residual Fuel Oil Imports |
| Energy Outlook | Residual Fuel Oil Product Supplied |
| Energy Outlook | Residual Fuel Oil Refinery and Blender Net Production |
| Energy Outlook | Residual Fuel OilStocks, End of Period |
| Energy Outlook | Solar/PV Energy Consumption/Production |
| Energy Outlook | Tight Oil Production |
| Energy Outlook | Total Biomass Energy Consumption |
| Energy Outlook | Total Biomass Energy Production |
| Energy Outlook | Total Energy Consumed by the Commercial Sector |
| Energy Outlook | Total Energy Consumed by the Industrial Sector |
| Energy Outlook | Total Energy Consumed by the Residential Sector |
| Energy Outlook | Total Energy Consumed by the Transportation Sector |
| Energy Outlook | Total Fossil Fuels Consumed by the Commercial Sector |

| | |
|---|---|
| Energy Outlook | Total Fossil Fuels Consumed by the Electric Power Sector |
| Energy Outlook | Total Fossil Fuels Consumed by the Industrial Sector |
| Energy Outlook | Total Fossil Fuels Consumed by the Residential Sector |
| Energy Outlook | Total Fossil Fuels Consumed by the Transportation Sector |
| Energy Outlook | Total Fossil Fuels Consumption |
| Energy Outlook | Total Petroleum Consumed by the Commercial Sector |
| Energy Outlook | Total Petroleum Consumed by the Electric Power Sector |
| Energy Outlook | Total Petroleum Consumed by the Industrial Sector |
| Energy Outlook | Total Petroleum Consumed by the Residential Sector |
| Energy Outlook | Total Petroleum Consumed by the Transportation Sector |
| Energy Outlook | Total Petroleum Imports |
| Energy Outlook | Total Petroleum Products Supplied |
| Energy Outlook | Total Petroleum Refinery and Blender Net Inputs |
| Energy Outlook | Total Petroleum Refinery and Blender Net Production |
| Energy Outlook | Total Primary Energy Consumption |
| Energy Outlook | Total Primary Energy Production |
| Energy Outlook | Total Renewable Energy Consumption |
| Energy Outlook | Total Renewable Energy Production |
| Energy Outlook | Unleaded Premium Gasoline, U.S. City Average Retail Price |
| Energy Outlook | Unleaded Regular Gasoline, U.S. City Average Retail Price |
| Energy Outlook | US Crude Oil First Purchase Price |
| Energy Outlook | US Crude Oil Stocks in Transit (on Ships) from Alaska |
| Energy Outlook | US less Alaskan North Slope Crude Oil First Purchase Price |
| Energy Outlook | Vehicle Miles Traveled, U.S. (million miles/day) |
| Energy Outlook | Waste Energy Consumption |
| Energy Outlook | Wind Energy Consumption/Production |
| Energy Outlook | Wood Energy Consumption |
| Energy Outlook | World Total Oil Net Imports |
| Financial and Commodity Markets | abrdn China A Share Equity R GOPRX |
| Financial and Commodity Markets | Brent Crack Spread |
| Financial and Commodity Markets | CBOE Crude Oil Volatility Index (OVX) |
| Financial and Commodity Markets | CBOE S&P 100 Volatility Index: VXO |
| Financial and Commodity Markets | CBOE SPX Volatility VIX |
| Financial and Commodity Markets | CRB BLS Spot Index (1967=100) |
| Financial and Commodity Markets | CRB BLS Spot Index Fats & Oils |
| Financial and Commodity Markets | CRB BLS Spot Index Foodstuffs |
| Financial and Commodity Markets | CRB BLS Spot Index Livestock |
| Financial and Commodity Markets | CRB BLS Spot Index Metals |
| Financial and Commodity Markets | CRB BLS Spot Index Raw Industrials |
| Financial and Commodity Markets | CRB BLS Spot Index Textiles |
| Financial and Commodity Markets | Crude Oil Non-Commercial Net Long Ratio |
| Financial and Commodity Markets | Cushing, OK Crude Oil Future Contract 1,2,3,4 |
| Financial and Commodity Markets | Dow Jones index |
| Financial and Commodity Markets | EURO STOXX 50 |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 1 month |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 10 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 11 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 12 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 2 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 24 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 3 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 36 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 4 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 48 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 5 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 6 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 60 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 7 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 72 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 8 months |
| Financial and Commodity Markets | Futures Brent crude oil, Intercontinental Exchange (ICE), 9 months |
| Financial and Commodity Markets | Hang Seng Index |
| Financial and Commodity Markets | MSCI Emerging Markets U$ |
| Financial and Commodity Markets | MSCI World U$ |
| Financial and Commodity Markets | NASDAQ Composite index |
| Financial and Commodity Markets | New York Harbor Conventional Gasoline Spot Price (GPNY) |
| Financial and Commodity Markets | New York Harbor No. 2 Heating Oil Future Contract 1 |
| Financial and Commodity Markets | New York Harbor No. 2 Heating Oil Future Contract 3 |
| Financial and Commodity Markets | NYSE AMEX Composite Index |
| Financial and Commodity Markets | NYSE Arca Oil Index |

| | |
|---|---|
| Financial and Commodity Markets | NYSE Composite Index |
| Financial and Commodity Markets | S&P GSCI Energy Total Return RETURN IND. (OFCL) |
| Financial and Commodity Markets | S&P 500 ES ENERGY |
| Financial and Commodity Markets | S&P 500 Index |
| Financial and Commodity Markets | S&P Goldman Sachs Commodity Index (GSCI) |
| Financial and Commodity Markets | S&P GSCI Non-Energy index |
| Financial and Commodity Markets | S&P' s Common Stock Price Index: Composite |
| Financial and Commodity Markets | S&P' s Composite Common Stock: Dividend Yield |
| Financial and Commodity Markets | S&P' s Common Stock Price Index: Industrials |
| Financial and Commodity Markets | S&P' s Composite Common Stock: Price-Earnings Ratio |
| Financial and Commodity Markets | Shanghai Composite Index |
| Financial and Commodity Markets | Stock Price of Major Oil Companies (XOM) |
| Financial and Commodity Markets | The ratio of trading volume of oil futures contracts to global oil production |
| Financial and Commodity Markets | Thomson Reuters Equal Weight Continuous Commodity Index (CCI) Energy 1967 = 100 |
| Financial and Commodity Markets | Thomson Reuters Equal Weight Continuous Commodity Index (CCI) Energy 1977=100 |
| Financial and Commodity Markets | Thomson Reuters Equal Weight Continuous Commodity Index (CCI) Grains & Oilseed |
| Financial and Commodity Markets | Thomson Reuters Equal Weight Continuous Commodity Index (CCI) Industrials |
| Financial and Commodity Markets | Thomson Reuters Equal Weight Continuous Commodity Index (CCI) Interest Rates |
| Financial and Commodity Markets | Thomson Reuters Equal Weight Continuous Commodity Index (CCI) Livestock Index |
| Financial and Commodity Markets | Thomson Reuters Equal Weight Continuous Commodity Index (CCI) Precious Metals |
| Financial and Commodity Markets | Thomson Reuters Equal Weight Continuous Commodity Index (CCI) Softs Index |
| Financial and Commodity Markets | US Gulf Coast Conventional Gasoline Spot Price (GPUS) |
| Financial and Commodity Markets | WTI Crack Spread |
| Financial and Commodity Markets | WTI Future-Spot Price Spread |
| Financial and Commodity Markets | WTI-Brent Spot Price Spread |
| Financial and Commodity Markets | Coal Price |
| Financial and Commodity Markets | Commodity Market Price Index |
| Financial and Commodity Markets | Copper Price |
| Financial and Commodity Markets | Gold Price |
| Financial and Commodity Markets | Iron Ore Price |
| Financial and Commodity Markets | Silver Price |
| Financial and Commodity Markets | Natural Gas Henry Hub Spot Price, U.S. (dollars per thousand cubic feet) |
| Financial and Commodity Markets | Bitcoin Price |
| Geopolitical Factors | Geopolitical Risk Index of Argentina |
| Geopolitical Factors | Geopolitical Risk Index of Brazil |
| Geopolitical Factors | Geopolitical Risk Index of China |
| Geopolitical Factors | Geopolitical Risk Index of Colombia |
| Geopolitical Factors | Geopolitical Risk Index of Hong Kong |
| Geopolitical Factors | Geopolitical Risk Index of India |
| Geopolitical Factors | Geopolitical Risk Index of Indonesia |
| Geopolitical Factors | Geopolitical Risk Index of Israel |
| Geopolitical Factors | Geopolitical Risk Index of Korea |
| Geopolitical Factors | Geopolitical Risk Index of Malaysia |
| Geopolitical Factors | Geopolitical Risk Index of Mexico |
| Geopolitical Factors | Geopolitical Risk Index of Philippines |
| Geopolitical Factors | Geopolitical Risk Index of Saudi Arabia |
| Geopolitical Factors | Geopolitical Risk Index of South Africa |
| Geopolitical Factors | Geopolitical Risk Index of Thailand |
| Geopolitical Factors | Geopolitical Risk Index of Turkey |
| Geopolitical Factors | Geopolitical Risk Index of Ukraine |
| Geopolitical Factors | Geopolitical Risk Index of Venezuela |
| Geopolitical Factors | Geopolitical Risk Index of World |
| Geopolitical Factors | Total amount of terrorist attack in the Middle East and North Africa |
| Housing | Housing Starts, Midwest |
| Housing | Housing Starts, Northeast |
| Housing | Housing Starts, South |
| Housing | Housing Starts, West |
| Housing | Housing Starts: Total New Privately Owned |
| Housing | New Private Housing Permits (SAAR) |
| Housing | New Private Housing Permits, Midwest (SAAR) |
| Housing | New Private Housing Permits, Northeast (SAAR) |
| Housing | New Private Housing Permits, South (SAAR) |
| Housing | New Private Housing Permits, West (SAAR) |
| Housing | S&P/Case-Shiller U.S. National Home Price Index |
| Industrial Production | Production of Total Industry in Austria |
| Industrial Production | Production of Total Industry in Belgium |
| Industrial Production | Production of Total Industry in Brazil |
| Industrial Production | Production of Total Industry in Canada |
| Industrial Production | Production of Total Industry in Czech Republic |
| Industrial Production | Production of Total Industry in Denmark |

| | |
|---|---|
| Industrial Production | Production of Total Industry in Finland |
| Industrial Production | Production of Total Industry in France |
| Industrial Production | Production of Total Industry in Germany |
| Industrial Production | Production of Total Industry in Greece |
| Industrial Production | Production of Total Industry in Hungary |
| Industrial Production | Production of Total Industry in Ireland |
| Industrial Production | Production of Total Industry in Israel |
| Industrial Production | Production of Total Industry in Italy |
| Industrial Production | Production of Total Industry in Japan |
| Industrial Production | Production of Total Industry in Korea |
| Industrial Production | Production of Total Industry in Netherlands |
| Industrial Production | Production of Total Industry in Norway |
| Industrial Production | Production of Total Industry in Poland |
| Industrial Production | Production of Total Industry in Portugal |
| Industrial Production | Production of Total Industry in Slovak Republic |
| Industrial Production | Production of Total Industry in Spain |
| Industrial Production | Production of Total Industry in Sweden |
| Industrial Production | Production of Total Industry in the United Kingdom |
| Industrial Production | Production of Total Industry in Turkey |
| Industrial Production | Industrial Production Index |
| Industrial Production | Industrial Production: Business Equipment |
| Industrial Production | Industrial Production: Consumer Goods |
| Industrial Production | Industrial Production: Durable Consumer Goods |
| Industrial Production | Industrial Production: Durable Materials |
| Industrial Production | Industrial Production: Construction supplies |
| Industrial Production | Industrial Production: Durable Goods: Alumina and aluminum production and processing |
| Industrial Production | Industrial Production: Durable Goods: Automotive products |
| Industrial Production | Industrial Production: Durable Goods: Cement and concrete product |
| Industrial Production | Industrial Production: Durable Goods: Iron and steel products |
| Industrial Production | Industrial Production: Durable Goods: Raw steel |
| Industrial Production | Industrial Production: Durable manufacturing: Aerospace and miscellaneous transportation |
| Industrial Production | Industrial Production: Durable manufacturing: Machinery |
| Industrial Production | Industrial Production: Durable manufacturing: Primary metal |
| Industrial Production | Industrial Production: Electric power generation, transmission, and distribution |
| Industrial Production | Industrial Production: Energy Materials: Energy, total |
| Industrial Production | Industrial Production: Final Products (Market Group) |
| Industrial Production | Industrial Production: Final Products and Nonindustrial Supplies |
| Industrial Production | Industrial Production: Fuels |
| Industrial Production | Industrial Production: Manufacturing(SIC) |
| Industrial Production | Industrial Production: Materials |
| Industrial Production | Industrial Production: Mining: Coal mining |
| Industrial Production | Industrial Production: Mining: Copper, nickel, lead, and zinc mining |
| Industrial Production | Industrial Production: Mining: Crude oil |
| Industrial Production | Industrial Production: Mining: Crude petroleum and natural gas extraction |
| Industrial Production | Industrial Production: Mining: Drilling oil and gas wells |
| Industrial Production | Industrial Production: Mining: Iron ore mining |
| Industrial Production | Industrial Production: Mining: Natural gas |
| Industrial Production | Industrial Production: Mining: Oil and gas extraction |
| Industrial Production | Industrial Production: Non energy, total |
| Industrial Production | Industrial Production: Nondurable Consumer Goods |
| Industrial Production | Industrial Production: Nondurable Goods: Chemical products |
| Industrial Production | Industrial Production: Nondurable Goods: Petroleum refineries |
| Industrial Production | Industrial Production: Nondurable Goods: Pharmaceutical and medicine |
| Industrial Production | Industrial Production: Nondurable Goods: Plastics material and resin |
| Industrial Production | Industrial Production: Nondurable manufacturing: Chemical |
| Industrial Production | Industrial Production: Nondurable manufacturing: Petroleum and coal products |
| Industrial Production | Industrial Production: Nondurable Materials |
| Industrial Production | Industrial Production: Nondurable manufacturing: Plastics and rubber products |
| Industrial Production | Industrial Production: Residential Utilities |
| Industrial Production | Capacity Utilization: Electric power generation, transmission, and distribution |
| Industrial Production | Capacity Utilization: Oil and gas extraction |
| Industrial Production | Capacity Utilization: Manufacturing |
| Interest and exchange rates | 1 Year Treasury C Minus FEDFUNDS |
| Interest and exchange rates | 1 Year Treasury Rate |
| Interest and exchange rates | 10 Year Treasury C Minus FEDFUNDS |
| Interest and exchange rates | 10 Year Treasury Rate |
| Interest and exchange rates | 3 Month AA Financial Commercial Paper Rate |
| Interest and exchange rates | 3 Month Commercial Paper Minus FED FUNDS |
| Interest and exchange rates | 3 Month Treasury Bill |
| Interest and exchange rates | 3 Month Treasury C Minus FEDFUNDS |

| | |
|---|---|
| Interest and exchange rates | 3-Month Treasury Constant Maturity Rate |
| Interest and exchange rates | 5 Year Treasury C Minus FEDFUNDS |
| Interest and exchange rates | 5 Year Treasury Rate |
| Interest and exchange rates | 6 Month Treasury Bill |
| Interest and exchange rates | 6 Month Treasury C Minus FEDFUNDS |
| Interest and exchange rates | Effective Federal Funds Rate |
| Interest and exchange rates | ICE US Dollar Index Futures Open Interest |
| Interest and exchange rates | LIBOR Rate |
| Interest and exchange rates | Moody' s Aaa Corporate Bond Minus FEDFUNDS |
| Interest and exchange rates | Moody' s Baa Corporate Bond Minus FEDFUNDS |
| Interest and exchange rates | Moody' s Seasoned Aaa Corporate Bond Yield |
| Interest and exchange rates | Moody' s Seasoned Baa Corporate Bond Yield |
| Interest and exchange rates | Trade Weighted U.S. Dollar Index |
| Interest and exchange rates | US 10-Year Bond Yield |
| Interest and exchange rates | US Default Yield Spread |
| Interest and exchange rates | US Real Interest Rate |
| Interest and exchange rates | US Short-term Interest Rate |
| Interest and exchange rates | US Term Spread |
| Interest and exchange rates | US Dollar index DXY |
| Interest and exchange rates | US/ Taiwan, Russia, Japan, U.K, Canada, Switzerland Foreign Exchange Rate |
| Labor market | All Employees: Construction |
| Labor market | All Employees: Durable goods |
| Labor market | All Employees: Financial Activities |
| Labor market | All Employees: Goods Producing Industries |
| Labor market | All Employees: Government |
| Labor market | All Employees: Manufacturing |
| Labor market | All Employees: Mining and Logging: Mining |
| Labor market | All Employees: Nondurable goods |
| Labor market | All Employees: Retail Trade |
| Labor market | All Employees: Service Providing Industries |
| Labor market | All Employees: Total nonfarm |
| Labor market | All Employees: Trade, Transportation & Utilities |
| Labor market | All Employees: Wholesale Trade |
| Labor market | Average Duration of Unemployment (Weeks) |
| Labor market | Avg Hourly Earnings : Construction |
| Labor market | Avg Hourly Earnings : Goods Producing |
| Labor market | Avg Hourly Earnings : Manufacturing |
| Labor market | Avg Weekly Hours : Goods Producing |
| Labor market | Avg Weekly Hours : Manufacturing |
| Labor market | Avg Weekly Overtime Hours : Manufacturing |
| Labor market | Civilian Employment |
| Labor market | Civilian Labor Force |
| Labor market | Civilian Unemployment Rate |
| Labor market | Civilians Unemployed 15 Weeks & Over |
| Labor market | Civilians Unemployed Less Than 5 Weeks |
| Labor market | Civilians Unemployed for 15 26 Weeks |
| Labor market | Civilians Unemployed for 27 Weeks and Over |
| Labor market | Civilians Unemployed for 5 14 Weeks |
| Labor market | Help Wanted Index for the United States |
| Labor market | Initial Claims |
| Labor market | Ratio of Help Wanted/No. Unemployed |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Australia |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Austria |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Belgium |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Big four European |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Brazil |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Canada |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Chile |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for China |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Denmark |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Euro area (19 countries) |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Finland |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for France |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for G7 |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Germany |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Greece |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Hungary |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Ireland |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Italy |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Japan |

| | |
|---|---|
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Korea |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Major five Asia |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Mexico |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for NAFTA |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Netherlands |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Norway |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for OECD Europe |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for OECD Major six NME |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for OECD Total |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Poland |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Portugal |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Russia |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for South_Africa |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Spain |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Sweden |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Switzerland |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for Turkey |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for United Kingdom |
| Macroeconomic indicators | OECD Composite Leading Indicator (CLI) for United States of America |
| Macroeconomic indicators | Chicago Fed National Activity Index |
| Macroeconomic indicators | China, US, Euro GDP Growth |
| Macroeconomic indicators | Kilian Global economic index |
| Macroeconomic indicators | World GDP Growth |
| Macroeconomic indicators | US Aruoba Diebold Scotti Business Conditions Index |
| Macroeconomic indicators | Leading Index for the United States |
| Macroeconomic indicators | NBER based Recession Indicators for the United States |
| Macroeconomic indicators | PMI Index |
| Macroeconomic indicators | University of Michigan: Inflation Expectation |
| Macroeconomic indicators | University of Michigan: Consumer Sentiment |
| Macroeconomic indicators | Real Personal Income |
| Macroeconomic indicators | Real personal income ex transfer receipts |
| Money and credit | Commercial and Industrial Loans |
| Money and credit | Consumer Motor Vehicle Loans Outstanding |
| Money and credit | M 1 Money Stock |
| Money and credit | M 2 Money Stock |
| Money and credit | Monetary Base |
| Money and credit | MZM Money Stock |
| Money and credit | Nonrevolving consumer credit to Personal Income |
| Money and credit | Real Estate Loans at All Commercial Banks |
| Money and credit | Real M 2 Money Stock |
| Money and credit | Reserves Of Depository Institutions |
| Money and credit | Securities in Bank Credit at All Commercial Banks |
| Money and credit | Total Consumer Loans and Leases Outstanding |
| Money and credit | Total Non revolving Credit |
| Money and credit | Total Reserves of Depository Institutions |
| Money and credit | Total Assets (US$ trillions), Federal Reserve |
| Money and credit | Total Assets (US$ trillions), Federal Reserve + European Central Bank + Bank of Japan |
| Prices | CPI : All Items Less Food |
| Prices | CPI : All items less medical care |
| Prices | CPI : All items less shelter |
| Prices | CPI : Apparel |
| Prices | CPI : Commodities |
| Prices | CPI : Durables |
| Prices | CPI : Medical Care |
| Prices | CPI : Services |
| Prices | CPI : Transportation |
| Prices | CRB Raw Materials Index |
| Prices | PPI: All Commodities |
| Prices | PPI: Crude Materials |
| Prices | PPI: Finished Consumer Goods |
| Prices | PPI: Finished Goods |
| Prices | PPI: Intermediate Materials |
| Prices | PPI: Metals and metal products |
| Prices | US CPI: All Items |
| Prices | US: CPI index: Energy |
| Prices | US: PPI index: Energy |
| Prices | US: PPI: manufacturing sector total |
| Prices | US: PPI: mining sector total |
| Prices | World CPI |