**Inter-Market Dynamics: Predicting Movement in the U.S. Dollar Index Based on Gold, Crude Oil, and Bitcoin Trends**



**Date: 10th December 2024**

**Team**

Appiah-Kubi Racheal (A20500031)

Dominguez Faraco Alfredo

Hemmati Hamed (A20460732)

Hosseini Mahla (A20499084)

Contents

[**ABSTRACT** 3](#_Toc183981940)

[**INTRODUCTION** 3](#_Toc183981941)

[**METHODOLOGY** 3](#_Toc183981942)

[**LITERATURE REVIEW** 4](#_Toc183981943)

[**BUSINESS UNDERSTANDING** 7](#_Toc183981944)

[Objective 7](#_Toc183981945)

[Business Questions 7](#_Toc183981946)

[**DATA UNDERSTANDING** 7](#_Toc183981947)

[Exploratory Data Analysis 7](#_Toc183981948)

[Correlation & Outlier Analysis 10](#_Toc183981949)

[Time Series Analysis 11](#_Toc183981950)

[**DATA PREPARATION** 14](#_Toc183981951)

[Feature Engineering 14](#_Toc183981952)

[Preprocessing & Pipelines 23](#_Toc183981953)

[**MODELLING & EVALUATION** 23](#_Toc183981954)

[Training 23](#_Toc183981955)

[Evaluation 23](#_Toc183981956)

[**DEPLOYMENT** 28](#_Toc183981957)

[Business Case 28](#_Toc183981958)

[Documentation 30](#_Toc183981959)

[Project Repository 30](#_Toc183981960)

[App 30](#_Toc183981961)

[**CONCLUSION & FUTURE WORK** 31](#_Toc183981962)

[**WHAT WE LEARNED** 31](#_Toc183981963)

[**REFERENCES** 31](#_Toc183981964)

[**APPENDIX & GLOSSARY** 32](#_Toc183981965)

# **ABSTRACT**

This project examines the inter-market dynamics among gold, crude oil, and Bitcoin and their collective influence on the U.S. Dollar Index (DXY). Using daily price data from 2014 to 2024, it employs the CRISP-DM methodology to uncover patterns and construct a predictive model for DXY movements. Key findings highlight the critical role of gold and crude oil volatility and percentage changes in forecasting DXY fluctuations. The study also proposes actionable portfolio strategies for risk management and return optimization, offering a foundation for data-driven investment decisions and scalable financial tools.

# **INTRODUCTION**

Global financial markets are interconnected, with price changes in one asset class often influencing others. This project investigates how gold, crude oil, and Bitcoin collectively impact the U.S. Dollar Index (DXY). By analyzing daily price data over a decade (2014–2024), the study aims to uncover dependencies among these assets and develop a predictive model for DXY. This model provides insights for portfolio risk evaluation and market forecasting, contributing to smarter investment strategies.

# **METHODOLOGY**

The Cross Industry Standard Process for Data Mining (CRISP-DM) framework was used as the base methodology for this project. It offers a structured approach, guiding us through key phases of *Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation,* and *Deployment.*

1. **Business Understanding:** Here we define the project goals, outlined business questions, and established the scope of analysis to understand the relationships between assets and their impact on the U.S. Dollar Index (DXY).
2. **Data Understanding:** Conducted an exploratory analysis of the collected dataset, examining time series trends, correlations, and identifying outliers to assess data quality and relevance.
3. **Data Preparation:** Applied preprocessing techniques, including cleaning and feature engineering, to prepare the data for modeling. Pipelines were established for consistent transformations and reproducibility.
4. **Modeling & Evaluation:** Built predictive models using historical data and evaluated their performance using appropriate metrics. The analysis included risk assessment to ensure practical utility for portfolio management.
5. **Deployment:** Developed a project repository and a user-friendly application to demonstrate the predictive model’s capabilities and its potential integration into risk management workflows.

# **LITERATURE REVIEW**

This project builds on existing research examining the relationships among financial assets, commodities, and the U.S. Dollar Index. These studies collectively provide a foundation for analyzing how gold, crude oil, and Bitcoin impact the U.S. Dollar Index, especially under volatile economic conditions. Key studies include:

**Inter-Market Relationships:**

1. *Attarzadeh* and *Balcilar* (2022) explored how Bitcoin, crude oil, clean energy, and stock markets interacted during crises, such as the COVID-19 pandemic, using a time-varying parameter model.

This study uses daily data from November 2013 to September 2021 to analyze the connectedness between Bitcoin, crude oil, clean energy, and stock markets. It employs a time-varying parameter model to examine how these markets interact, particularly during crises like the 2018 cryptocurrency crash and the COVID-19 pandemic in 2020. Both this study and our project focus on how Bitcoin and crude oil interact with other markets. While they explore asset connectedness, our project specifically investigates how Bitcoin, crude oil, and gold influence the U.S. Dollar Index (DXY).

1. *Zhang* and *Zhou* (2023) analyzed the U.S. Dollar, crude oil, gold, and Bitcoin, focusing on higher-order moments like volatility and skewness, using Granger causality to reveal interconnections.

This study uses high-frequency data from January 2016 to June 2022 on the U.S. Dollar, crude oil, gold, and Bitcoin. The authors apply robust estimators for volatility, skewness, and kurtosis to analyze the higher-order moments of these assets, using Granger causality and impulse response analysis to assess cross-market interconnections. Both studies focus on the interactions between the U.S. Dollar, crude oil, gold, and Bitcoin. While this paper looks at higher-order moments and causality, our project aims to predict the U.S. Dollar Index using historical price data for these assets.

1. *Koczar* et al. (2024) studied dynamic linkages between crude oil, renewable energy indices, and U.S. market sectors, highlighting contagion effects during events like the COVID-19 crisis.

This study uses daily data from August 2014 to February 2024 to assess the dynamic relationships between crude oil futures contracts, renewable energy indices, carbon credit futures indices, and various U.S. sector market indices. The authors apply the DECO-GARCH model and the Diebold-Yilmaz connectedness index to analyze these relationships, focusing on periods of stress, such as the COVID-19 crisis and the Russian Ukrainian conflict. Both this study and our project explore how crude oil prices influence other markets. While this paper focuses on oil’s interconnectedness with various U.S. sectors and energy markets, our project focuses on how oil prices, along with gold and Bitcoin, affect the U.S. Dollar Index.

**Bitcoin and Gold as Financial Tools:**

1. *Baur* and *McDermott* (2009) investigated gold’s role as a safe haven during market instability.

This study analyzes a 30-year period from 1979 to 2009, examining the role of gold as a hedge and safe haven across major global stock markets. It investigates how gold behaves compared to stocks in various countries, particularly during periods of market instability, including the financial crisis. Both this study and our project focus on gold’s role as a financial asset, particularly in times of economic stress. While this paper looks at gold’s safe haven behavior against stock market losses, our project examines how gold impacts the U.S. Dollar Index in combination with other assets like crude oil and Bitcoin.

1. *Baur* and *Hoang* (2021) examined Bitcoin's correlation with gold, showing a fluctuating relationship depending on market conditions.

This study uses daily and monthly price data from 2011 to 2021 for Bitcoin and gold to explore their correlation. It finds a near-zero correlation Project Proposal - Inter-Market Dynamics: Predicting the U.S. Dollar Index Based on Gold, Crude Oil, and Bitcoin Trends between Bitcoin and gold, challenging the idea of Bitcoin as "digital gold." The authors suggest the correlation fluctuates depending on market conditions, such as safe-haven events. This study is relevant to our analysis of Bitcoin and gold. However, while it focuses on the correlation between Bitcoin and gold, our project broadens the scope by including crude oil and investigating how all three assets collectively influence the U.S. Dollar Index.

**Oil Prices and Currency Dynamics:**

1. *Das* and *Jana* (2020) compared Bitcoin, gold, and the U.S. Dollar as hedges against oil price volatility.

This study uses daily data on Bitcoin, crude oil volatility (OVX), gold, commodities, and the U.S. Dollar. It looks at how Bitcoin performs as a hedge compared to gold and the U.S. Dollar, especially during times of oil price changes. Both studies explore how Bitcoin, gold, and crude oil interact with other markets. While this study looks at Bitcoin as a hedge, our project investigates how these assets influence the U.S. Dollar Index.

1. *Lin* et al. (2015) analyzed the effect of macroeconomic factors like U.S. monetary policy on the dynamic relationships between the dollar, oil, and gold prices.

This study uses weekly data from 1982 to 2013 on the U.S. Dollar, crude oil, and gold prices, sourced from DataStream. It applies wavelet analysis to explore the dynamic relationships between these assets, with a focus on how macroeconomic conditions, such as U.S. monetary policy, affect these correlations, particularly during financial crises. Both this study and our project investigate the relationship between the U.S. Dollar, crude oil, and gold. While they focus on oil-gold-dollar interactions, we extend this analysis by including cryptocurrencies like Bitcoin and examining their influence on the U.S. Dollar Index.

1. *Suliman* and *Abid* (2020) confirmed a long-term relationship between oil prices and exchange rates in Saudi Arabia.

This study uses monthly data from January 1986 to March 2019 to explore the relationship between oil prices and exchange rates in Saudi Arabia. The authors apply an autoregressive distributed lag model and an error correction model to identify long-term and short-term dynamics, confirming a robust long-term cointegration between oil prices and exchange rates. Both this study and our project examine the relationship between oil prices and exchange rates. However, while this study focuses on the Saudi Arabian exchange rate, our project expands the analysis to include the U.S. Dollar Index and how oil prices, along with gold and Bitcoin, influence it.

**Cryptocurrency Volatility:**

1. *Bakas* and *Magkonis* (2022) explored drivers of Bitcoin market volatility, emphasizing the influence of external shocks.

This study uses monthly data from January 1997 to July 2013 to analyze the impact of oil price shocks on stock market performance in three distinct Asian economies. The authors apply Toda and Yamamoto’s causality approach and Project Proposal - Inter-Market Dynamics: Predicting the U.S. Dollar Index Based on Gold, Crude Oil, and Bitcoin Trends generalized impulse response functions to examine both long-term and short-term relationships between oil prices and stock market returns. Both studies investigate the influence of oil prices on financial markets. While this paper focuses on the effect of oil price shocks on stock markets in different economies, our project explores how oil prices, along with gold and Bitcoin, impact the U.S. Dollar Index.

# **BUSINESS UNDERSTANDING**

## Objective

The primary objective of this project is to predict fluctuations in the U.S. Dollar Index (DXY) using historical prices of Bitcoin, crude oil, and gold, alongside other influential factors such as economic trends and temporal variables. Additionally, the project aims to utilize these predictions to ***calculate*** and ***assess portfolio risk***, providing actionable insights for more informed investment strategies.

## Business Questions

1. What are the relationships between Bitcoin, crude oil, and gold prices, and how do they correlate with changes in the U.S. Dollar Index?
2. To what extent do external features, such as holidays, weekdays, or the DXY itself, influence the accuracy and predictive power of the model?
3. How can the predictive model be used to calculate and evaluate portfolio risk, helping investors optimize their asset allocation?
4. What actionable insights can the predictive model offer for market forecasting and portfolio management strategies?

# **DATA UNDERSTANDING**

We began with individual data sets for all four commodities: ***gold***, ***crude oil***, ***bitcoin***, and the ***U.S. Dollar Index (DXY)***. Each data set spans from 1st November 2014 to 31st October 2024, marking a 10-year period. This comprehensive timeframe allows us to capture long-term market trends, seasonal variations, and economic cycles that impact these financial assets, setting the stage for in-depth analysis and predictive modeling.

## Exploratory Data Analysis

The dataset consisted of 3,653 rows and 23 columns, capturing detailed information on commodity prices and related features. Key features included the following:

* **Bitcoin (BTC):** Opening, high, low, and closing prices for Bitcoin.
* **Oil:** Opening, high, low, and closing prices for crude oil.
* **Gold:** Opening, high, low, and closing prices for gold.
* **DXY:** Opening, high, low, and closing values for the U.S. Dollar Index (DXY).
* **Date:** The date when the commodity values were recorded.
* **Temporal Variables:**
  + **Year:** The calendar year of the data point.
  + **Month:** The calendar month of the data point.
* **Holiday Indicators:**
  + **is\_holiday:** A binary indicator denoting whether the date was a holiday.
  + **Title:** The specific name of the holiday, if applicable.
* **is\_Weekday:** A binary indicator specifying whether the date fell on a weekday.

**Distribution Analysis**

These histograms illustrate the distributions of the opening, high, low, and closing prices for **Bitcoin (BTC),** **crude oil**, **gold**, and **dxy** (dollar price index). Here are the key takeaways:

**Bitcoin (BTC)**

A graph of different sizes and numbers

Description automatically generated with medium confidence

The distributions of Bitcoin's open, high, low, and close prices are heavily right skewed. Most prices are concentrated in the lower range, with only a few instances of higher values, indicating periods of significant price spikes.

**Crude Oil**

A graph of a number of numbers

Description automatically generated with medium confidence

The distributions of crude oil prices (open, high, low, and close) are more symmetrical compared to Bitcoin. The prices are primarily concentrated in the $25–$75 range, suggesting ***less extreme volatility*** compared to Bitcoin.

**Gold**

**A graph of a graph

Description automatically generated with medium confidence**

The distributions for gold prices (open, high, low, close) are slightly skewed. Most values fall between **$1,000 and $2,000**, indicating a relatively narrow range for the majority of observations. There are fewer occurrences of higher prices beyond $2,000, suggesting less frequent extreme price spikes.

**DXY (U.S. Dollar Index)**

**A graph of a graph

Description automatically generated with medium confidence**

The distributions of DXY prices (open, high, low, close) are roughly symmetrical. Most values are concentrated between **90 and 110**, indicating a stable price range for the index over the observed period. This stability suggests that the DXY exhibits much lower volatility compared to gold or Bitcoin.

## Correlation & Outlier Analysis

Bitcoin and gold are highly correlated, making them more likely to move together. The U.S. Dollar Index has weaker correlations with bitcoin, but much stronger correlations with oil and gold at approximately 0.44.

A graph of a heatmap

Description automatically generated

Outlier Analysis

A screenshot of a graph

Description automatically generated

**Oil** prices exhibit greater variability and more extreme outliers, highlighting the influence of market shocks and geopolitical events.

**DXY** is more stable, with fewer and less extreme outliers, reflecting its function as a global currency benchmark.

**Bitcoin** exhibits high volatility and extreme outliers, reflecting its speculative and unpredictable nature.

**Gold** demonstrates stability with fewer and less extreme outliers, reinforcing its role as a safe-haven asset in financial markets.

## Time Series Analysis

*Bitcoin & Crude Oil*

A graph showing the growth of the stock market

Description automatically generated

**Bitcoin (BTC)**

* Bitcoin's price exhibits exponential growth and extreme volatility over the observed period.
* Significant price increases are observed around **2017 (initial surge)** and **2020–2021 (another major spike)** and the price stabilizes at higher levels toward 2024–2025.

**Crude Oil**

* Crude oil prices show moderate fluctuations compared to Bitcoin, reflecting a more stable and cyclical pattern.
* A sharp decline around **2020**, likely linked to the COVID-19 pandemic and reduced global demand. Prices gradually decline and stabilize from **2023 onward**.

*Gold & DXY*

A graph showing the growth of the stock market

Description automatically generated

**Gold**

* Gold prices display a steady upward trend over the observed period, reflecting its role as a safe-haven asset.
* Notable spikes occur around **2020** (potentially linked to the COVID-19 pandemic) and **2024–2025**, indicating increased demand for gold during times of economic uncertainty.
* The steep rise in prices toward 2025 suggests heightened market instability or geopolitical events driving gold's value higher.

**DXY (U.S. Dollar Index)**

* DXY exhibits cyclical fluctuations, with peaks around **2017** and **2022**, followed by declines and subsequent recoveries.
* A sharp dip is observed around **2020**, likely tied to the pandemic’s impact on the global economy.

**Decomposition**

Decomposition is a statistical method used in time series analysis to break down a series into its key components. This helps in understanding the underlying patterns and behaviors within the data.

We will look at the seasonal decomposition of our predictor feature **Dollar Price Index.**

A graph of a graph of a graph

Description automatically generated with medium confidence

**Observed (dxy\_close):** The top plot shows the original time series, which is the ***dxy\_close*** values over time. It displays the overall behavior of the series, including trends, seasonality, and irregular components.

**Trend:** The second plot represents the trend component. It reflects the long-term movement or direction in the data, excluding short-term fluctuations and seasonal patterns. In this case, there is a clear upward trend from 2020 to 2022, followed by a slight decline in 2023.

**Seasonal:** The third plot captures the repeating seasonal patterns in the data. This component highlights periodic fluctuations, which could be driven by recurring factors, such as market cycles or seasonal economic behavior.

Here, the seasonal component is regular and stable over the years, showing periodic ups and downs that recur consistently.

**Residual (Noise):** The bottom plot shows the residuals, which are the unexplained components after removing the trend and seasonal effects. These are random fluctuations or noise in the data. The residuals seem more volatile around 2020 to the end of 2023, and this could be attributed to the COVID-19 Vaccine.

*NB: Find seasonal decomposition for the 3 other commodities in the appendix.*

# **DATA PREPARATION**

## Feature Engineering

In the complex landscape of financial markets, effective feature engineering is crucial for distilling raw data into insights. This section outlines our methodology for developing key features tailored to enhance the model. Each subset of engineered features is meticulously described, detailing the transformation processes applied and their relevance to the overall modeling effort. These descriptions are complemented by visualizations that demonstrate the behavioral patterns of each feature, providing a clearer understanding of their dynamics and potential utility in predictive modeling. While not all features may survive the rigors of feature selection or overcome challenges like multicollinearity, their inception was grounded in comprehensive exploratory data analysis.

To streamline this documentation, the placeholder {asset} is used throughout this section to denote features that are applicable across four major financial indicators: Bitcoin (BTC), gold, oil, and the Dollar Index (DXY). This notation facilitates a concise and consistent presentation of feature engineering efforts across different assets.

* **Monthly and weekly Rolling averages**

Features’ Name: {asset}\_weekly\_avg, {asset}\_monthly\_avg

Description: Weekly and monthly moving averages of {asset}\_close to smooth out daily fluctuations and reveal longer-term trends.

Transformation Logic: Calculated as the rolling means over a 7-day and 30-day window.

Datatype and missing values: Numeric, first value calculated after the first complete window.

Visualization: The plot illustrates how rolling averages effectively smoothed out the volatility in daily prices, highlighting the underlying trend.

**A graph of a stock market

Description automatically generated with medium confidence**

* **Percentage changes in Monthly and weekly Rolling averages**

Features’ Name: {asset}\_weekly\_avg\_pct\_change, {asset}\_monthly\_avg\_pct\_change

Description: Measures the rate of change in the weekly/monthly moving averages, indicating the strength and direction of the trend.

Transformation Logic: Percentage change between consecutive {asset}\_weekly\_avg, {asset}\_monthly\_avg.

Datatype and missing values: Numeric, first value calculated after the first complete window.

Visualization: The plots depict the frequency of different percentage changes in weekly and monthly average prices for each asset. The narrower distribution of monthly average price changes indicates a lower degree of fluctuation over longer time periods.

A graph of different changes

Description automatically generated with medium confidence

* **Percentage changes in Monthly and weekly Rolling averages**

Features’ Name: {asset}\_Daily\_ocpercentage, {asset}\_Daily\_max\_percentage, {asset}\_Daily\_min\_percentage

Description: Percentage changes in daily open, close, high, and low prices to quantify the magnitude and frequency of price swings.

Transformation Logic: {asset}\_open to {asset}\_close, {asset}\_close to {asset}\_high, and {asset}\_close to {asset}\_low percentage changes.

Datatype and missing values: Numeric, No missing.

Visualization: The histograms visualize the distribution of the new features for each asset. The plots reveal that Bitcoin and Oil exhibit higher volatility compared to Gold and the Dollar Index, with wider distributions and more frequent extreme price movements.

A graph of different colored lines

Description automatically generated with medium confidence

* **Percentage changes in Monthly and weekly Rolling averages**

Feature’s Name: {asset}\_intraday\_volatility

Description: Calculates the percentage difference between the daily high and low prices, providing a measure of the magnitude of price swings within a single trading day.

Transformation Logic: Calculated as the percentage difference between the {asset}\_high and {asset}\_low divided by the {asset}\_close.

Datatype and missing values: Numeric, No missing.

Visualization: The plots show that Bitcoin and Oil exhibit significantly higher intraday volatility compared to Gold and the Dollar Index, with wider distributions and more frequent occurrences of extreme price swings within a single trading day.

A group of graphs showing different types of volatility

Description automatically generated

* **Day to day Percentage changes in prices (% Return)**

Feature’s Name: {asset}\_daily\_percentage

Description: Percentage change in the daily closing price to quantify the daily performance and volatility of the asset.

Transformation Logic: Percentage difference between today’s {asset}\_close and yesterday’s {asset}\_close.

Datatype and missing values: Numeric, Forward fill for the first day's missing value.

Visualization: The plots show that Bitcoin and Oil exhibit significantly higher intraday volatility compared to Gold and the Dollar Index, with wider distributions and more frequent occurrences of extreme price swings within a single trading day.

A graph of different types of graphs

Description automatically generated with medium confidence

* **Weekly and Monthly Volatilities**

Features’ Name: {asset}\_rolling\_volatility\_7, {asset}\_rolling\_volatility\_30

Description: 7-day and 30-day rolling volatility measures the fluctuations in daily price changes over specific time windows. This metric helps gauge short-term and medium-term market risk.

Transformation Logic: Standard deviation of {asset}\_daily\_percentage over 7-day and 30-day windows.

Datatype and missing values: Numeric, Forward fill for the first day's missing value.

Visualization: The plots show that Bitcoin and Oil exhibit significantly higher volatility compared to Gold and the Dollar Index, with wider distributions and more frequent occurrences of extreme volatility periods.

A graph of different colored lines

Description automatically generated with medium confidence

* **Interaction Features: Ratios**

Features’ Name: gold\_oil\_ratio, gold\_btc\_ratio, oil\_btc\_ratio

Description: These features calculate the ratios between the closing prices of gold, oil, and Bitcoin. These ratios help to understand the relative strength and weakness of each asset compared to the others.

Transformation Logic: Division of the respective {asset}\_close.

Datatype and missing values: Numeric, Forward fill to handle any sporadic missing values in daily data.

Visualization: The plots show that the relationships are not linear, indicating potential non-linear interactions between these assets. This justifies the use of ratio features to capture these complex relationships and improve model performance.

A group of graphs showing different colored dots

Description automatically generated with medium confidence

**Interaction Features: Differences**

**Features’ Name:** dxy\_gold\_pct\_diff, dxy\_oil\_pct\_diff, dxy\_btc\_pct\_diff

Description: These features calculate the daily percentage difference between the Dollar Index (DXY) and other assets (Gold, Oil, and Bitcoin).

**Transformation Logic:** Differences between dxy\_daily\_percentage and other {asset}\_daily\_percentage to capture their relative performance each day.

**Datatype and missing values:** Numeric, No Missing.

**Visualization:** The line plots visualize the daily percentage changes of Gold, Bitcoin, and Oil relative to DXY over time. These plots highlight periods of relative strength and weakness for each asset compared to DXY, suggesting that these difference features can capture valuable information for the relationship between these commodities and DXY.

A screenshot of a graph

Description automatically generated

**Interaction Features: Differences**

Features’ Name: {asset}\_yesterday\_Daily\_ocpercentage,   
{asset}\_yesterday\_intraday\_volatility,   
{asset}\_yesterday\_daily\_percentage,   
{asset}\_yesterday\_weekly\_avg\_pct\_change,   
{asset}\_yesterday\_monthly\_avg\_pct\_change

Description: Each feature is shifted by one day to capture the previous day's value.

Transformation Logic: Differences between dxy\_daily\_percentage and other {asset}\_daily\_percentage to capture their relative performance each day.

Datatype and missing values: Numeric, Missing values for the first entry are filled with the next available value.

## Preprocessing & Pipelines

# **MODELLING & EVALUATION**

## Training

## Evaluation

The analysis of feature importance in predictive modeling offers a window into understanding which variables most influence the model's predictions. For the models which had the best performance in this study—Random Forest and Stacking Classifier—we extracted the effect of each feature on the model.

**Random Forest**

* **Feature importance Mechanism:** In the Random Forest model, the effect of each feature on the model's predictions is quantified using the .feature\_importances\_ method, which calculates the importance by looking at how much each feature decreases the impurity of a split in the trees that make up the forest. This measure is computed during the training phase and represents the average decrease in impurity caused by the splits that include the feature across all trees.

A graph of a number of different colored bars

Description automatically generated with medium confidence

* **High Importance:** The Random Forest model exhibits a pronounced preference for features associated with gold and oil percentage changes and volatilities, such as gold\_daily\_percentage, oil\_daily\_percentage, and their intraday volatilities. These features appear to capture significant predictive power, likely due to their direct impact on market behaviors and price movements.
* **Low Importance**: Time-related features like Month, Year, and is\_holiday show minimal influence on predictions. This suggests that the model, which likely focuses on shorter-term fluctuations rather than seasonal or annual trends, finds limited predictive value in these broader time indicators.

**Stacking Classifier:**

* **Feature importance Mechanism:** For the Stacking Classifier, the method to determine feature importance is not directly inherent but can be deduced from the base learners that make up the stack. Initially, we extract the feature importance of each of the base learners. Each base learner's importance is considered in the context of its contribution to the final prediction. Then, We also extract the meta-learner coefficients, which are then averaged across all classes if the model is a multi-class classifier. These averaged coefficients represent the weight each base learner contributes to the final decision. The overall feature importances for the Stacking Classifier are then calculated by multiplying each base learner's feature importances by their respective weights from the meta-learner and summing them up. The results are normalized (dividing by the mean) to provide a scale for comparison, producing a comprehensive list of features ranked by their weighted importance in the stacked model.

A graph of a number of different colored bars

Description automatically generated with medium confidence

* **High Importance**: The Stacking Classifier similarly emphasizes the importance of daily percentage changes and volatility, especially for gold and oil. Features like gold\_daily\_percentage and oil\_intraday\_volatility dominate, highlighting their relevance in predicting market dynamics.
* Both models underline the reduced importance of date-related features. Their negligible influence aligns with the focus on more immediate market indicators that offer real-time predictive value.
* The overlap in high-importance features between the two models reinforces the notion that certain market dynamics, particularly those related to gold and oil, are critical for prediction.

A graph with numbers and dots

Description automatically generated

**Feature Analysis by Asset Category**

The bar chart illustrates the relative importance of features associated with Gold, Oil, and Bitcoin, as assessed by Random Forest and Stacking Classifier models. Notably, both models consistently rank Gold features as most influential, with the Stacking Classifier attributing slightly greater significance (40.01%) compared to the Random Forest (36.11%). This suggests that fluctuations in Gold prices may have a substantial impact on DXY price movements, followed by Oil and Bitcoin.

A graph of different colored bars

Description automatically generated

**Lagged Feature Impact Analysis**

The bar chart illustrates the relative importance of lagged features for both the Random Forest and Stacking Classifier models. Both models exhibit a similar reliance on lagged features, attributing approximately 40% of their predictive power to these historical data points. This suggests that both models leverage past information to inform future predictions, highlighting the significance of historical context in your analysis. The consistency across these diverse models underscores the robustness of lagged features as key indicators in your predictive framework.

**A screenshot of a graph

Description automatically generated**

**Impact of Volatility Features on Model Performance**

The bar chart illustrates the relative importance of volatility-related features for both the Random Forest and Stacking Classifier models. Random Forest attributes 41.81% of its predictive power to volatility features, slightly higher than the Stacking Classifier's 40.36%. This indicates that both models heavily rely on volatility measures, showcasing their role in capturing dynamic market movements. Similar dependency across models suggests the robustness of volatility-related features in explaining and predicting variations within the dataset.

A graph of a graph with numbers and text

Description automatically generated with medium confidence

In conclusion, the evaluation of feature importance across the Random Forest and Stacking Classifier models highlights the critical role of certain market indicators, particularly gold and oil-related features, in predicting DXY price movements. Both models consistently emphasize the importance of daily percentage changes, volatility measures, and lagged features, reflecting their reliance on dynamic and historical market behaviors. Conversely, date-related features such as Month, Year, and is\_holiday were deemed less impactful, underscoring the models' focus on short-term and real-time predictors over broader seasonal trends. The overlap in high-importance features across models reinforces the robustness of these indicators, with volatility and lagged features emerging as pivotal components in the predictive framework. This alignment between models not only validates the choice of features but also enhances confidence in the reliability of their insights for market analysis.

# **DEPLOYMENT**

## Business Case

***Building Smarter Investment Portfolios using Movement in DXY***

The U.S. Dollar Index (DXY) serves as an essential indicator of the overall strength of the U.S. Dollar against a basket of major currencies, and its movements can provide significant insights into broader economic trends. A strengthening USD typically indicates growing confidence in the U.S. economy, often driven by factors such as rising interest rates, increased global demand for dollars, or safe haven buying during times of market turmoil. Conversely, a weakening USD suggests reduced confidence, which may arise from concerns about inflation, lower interest rates, or a shift in investor focus on riskier assets like stocks or commodities.

For investors, understanding the direction of the DXY is crucial, as it can inform decisions regarding the reallocation of funds into different assets that tend to perform well under varying market conditions. By analyzing historical data from assets like Bitcoin, Gold, and Oil, the model identifies vital patterns and relationships influencing DXY. Gold, for example, has traditionally been viewed as a hedge against a weak dollar; when the USD weakens, gold prices often rise as investors seek safer investment. Similarly, Oil, being priced in USD globally, can experience significant demand and price fluctuations influenced by changes in the dollar's value. On the other hand, Bitcoin, recognized for its volatility, often acts as a hedge during periods of economic uncertainty or instability in the USD.

Through careful analysis of how these different assets interact with the DXY under various conditions, such as market volatility and price changes, the model is capable of predicting whether the USD will strengthen, weaken, or remain stable. This predictive capability ultimately empowers investors to make informed decisions about their investment strategies in response to evolving market circumstances.

**Investment strategies**

Once the model predicts the movement of the DXY (U.S. Dollar Index), the focus shifts to portfolio allocation. Two distinct strategies are proposed based on the prediction:

1. **Minimizing risk:** When the model predicts a weakening USD, it signals potential market uncertainty. In this scenario, the strategy prioritizes lower-risk investments:

* During a period of a weak USD, riskier assets may become more volatile. To protect the investor, the portfolio shifts toward more stable options, such as Gold, which has historically acted as a safe haven during economic downturns.
* The asset weights are determined using the “inverse volatility” approach. This means that assets with lower price swings (lower volatility) receive a higher portion of the investment.

1. **Maximizing returns:** When the model predicts a strengthening USD, it indicates favorable conditions for the dollar, often associated with economic growth or increased global confidence. This creates opportunities to take on more risk for higher returns:

* With a strengthening USD, investors may capitalize on assets that perform well during periods of optimism or stability, such as Bitcoin, which offers higher potential returns despite its volatility.
* The asset weights are based on “Sharpe ratios,” which measure the return an asset provides for each unit of risk. Assets with higher Sharpe ratios receive a larger allocation, reflecting their superior risk-adjusted performance.

#### **Practical Examples**

#### **Scenario 1: Minimizing Risk**

* **Date**: October 1, 2020
* **Prediction**: Negative (USD Weakens)
* **Strategy**: Allocate funds to prioritize stability and reduce exposure to volatile assets.
* **Outcome**:
  + **Gold**: 64.17% ($577.54 of a $900 investment)
  + **Bitcoin**: 22.16% ($199.41)
  + **Oil**: 13.67% ($123.05)

#### **Scenario 2: Maximizing Returns**

* **Date**: January 3, 2024
* **Prediction**: Positive (USD Strengthens)
* **Strategy**: Allocate funds to assets with higher growth potential, accepting higher risk for greater rewards.
* **Outcome**:
  + **Bitcoin**: 69.24% ($623.19 of a $900 investment)
  + **Gold**: 19.25% ($173.29)
  + **Oil**: 11.50% ($103.51)

#### **Business Value: Why This Matters**

1. **A Reliable Guide for Investors:** Financial decisions are often clouded by uncertainty. This approach provides a structured, data-driven framework to inform investment decisions.

* **Predictive Insight**: The model forecasts key market movements, giving investors an early signal to adjust strategies.
* **Tailored Portfolios**: Recommendations are aligned with specific market conditions, catering to both risk-averse and risk-tolerant investors.

1. **Improved Investment Outcomes:** By aligning portfolio strategies with market predictions, investors can:

* **Minimize losses** during turbulent periods by leaning on safer assets.
* **Maximize gains** during favorable conditions by seizing high-return opportunities.

1. **Versatility Across Market Cycles:** The dual-strategy system ensures relevance in both bullish and bearish markets, helping investors stay ahead of market trends regardless of economic conditions.
2. **Simplified Decision-Making:** Investors no longer need to interpret complex market data. The model’s outputs are clear: a prediction of DXY’s movement and a corresponding portfolio recommendation.

This project connects complex market analysis with actionable investment strategies. By predicting movements in the US Dollar and aligning portfolios with these predictions, investors are equipped to make smarter and more confident decisions. This method not only improves individual investment results but also establishes a foundation for a scalable financial advisory tool that can adapt to changing markets.

## Documentation

### Project Repository

[Link to project repository on GitHub](https://github.com/coderacheal/Inter-Market-Dynamics-Gold-Oil-Bitcoin-Dollar-Index-)

### App

[Link to deployed App on Render](https://inter-market-dynamics-gold-oil-bitcoin.onrender.com)

# **CONCLUSION & FUTURE WORK**

# **WHAT WE LEARNED**

Through this project, we had the opportunity to explore and apply a range of new concepts and tools, including:

* **Base Models**: Developing and fine-tuning foundational predictive models for financial market analysis.
* **Financial Markets and Risk Calculation**: Gaining insights into inter-market dynamics and employing risk assessment techniques to evaluate portfolio strategies.
* **Streamlit**: Building an interactive and user-friendly application to showcase predictive capabilities.
* **GitHub**: Utilizing version control and collaboration tools for efficient project management and documentation.

# **REFERENCES**

1. Attarzadeh, A., & Balcilar, M. (2022). Dynamic return and volatility connectedness of cryptocurrency, crude oil, clean energy, and stock markets. *Science of the Total Environment, 827*, 154249. <https://doi.org/10.1016/j.scitotenv.2022.154249>
2. Baur, D. G., & Hoang, L. T. (2021). The Bitcoin-gold correlation puzzle. *Finance Research Letters, 38*, 101428. <https://doi.org/10.1016/j.frl.2020.101428>
3. Baur, D. G., & McDermott, T. K. J. (2009). Is gold a safe haven? International evidence. *Journal of Banking & Finance, 34*(8), 1886-1898. <https://doi.org/10.1016/j.jbankfin.2009.12.008>
4. Das, A., & Jana, P. (2020). Bitcoin as a hedge against oil price volatility: A comparison with gold, commodity, and the US Dollar. *International Journal of Financial Markets and Derivatives, 7*(4), 257–275. <https://doi.org/10.1504/IJFMD.2020.10027321>
5. Koczar, M. W., Jareño, F., & Escribano, A. (2024). Dynamic linkages and contagion effects between crude oil prices, US market sector indices, and energy markets. *Energy Economics, 118*, 106454. <https://doi.org/10.1016/j.eneco.2024.106454>
6. Lin, C., & Yang, S. (2015). Does the value of the US dollar matter with the price of oil and gold? *Energy Economics, 50*, 130–141. <https://doi.org/10.1016/j.eneco.2015.04.012>
7. Suliman, T. H. M., & Abid, M. (2020). The impacts of oil price on exchange rates: Evidence from Saudi Arabia. *International Economics, 163*, 174–190. <https://doi.org/10.1016/j.inteco.2020.03.005>
8. Zhang, Y., & Zhou, L. (2023). Higher-order moment nexus between the US Dollar, crude oil, gold, and Bitcoin. *North American Journal of Economics and Finance, 67*, 101694. <https://doi.org/10.1016/j.najef.2023.101694>
9. Chen, J. (2024, October 19). *What is the U.S. Dollar Index (USDX) and how to trade it*. Investopedia. <https://www.investopedia.com/terms/u/usdx.asp>
10. <https://www.kaspersky.com/resource-center/definitions/what-is-cryptocurrency>

# **APPENDIX & GLOSSARY**

* ***US Dollar Index*** - The U.S. dollar index is a measure of the value of the U.S. dollar relative to a basket of foreign currencies. The USDX was established by the U.S. Federal Reserve in 1973 after the dissolution of the Bretton Woods Agreement. It is now maintained by ICE Data Indices, a subsidiary of the Intercontinental Exchange (ICE).
* **Cryptocurrency** - Cryptocurrency, sometimes called crypto-currency or crypto, is any form of currency that exists digitally or virtually and uses cryptography to secure transactions.
* ***Decomposition*** is a statistical method used in time series analysis to break down a series into its key components (trend, seasonality, residuals, observed values)

A graph of oil prices

Description automatically generatedA graph of a graph of a graph

Description automatically generated with medium confidence

A graph of a graph of a graph

Description automatically generated with medium confidence