Forecasting Methods in Fintech
096292

Spring 2024

Final Project Exploration Phase

Github

Submitted by:

Bashar Massad
Lana Haj
Kamilya Ziyan
Majd Bishara
Sami Mor

Introduction

In this project, we are going to predict the price of Gold Futures contracts listed on the COMEX using multiple machine learning (ML) and time series models. Gold Futures reflect the price for one troy ounce of gold in U.S. dollars and are influenced by various economic and market factors. We collected daily data on features such as crude oil prices, silver prices, stock indices, and currency exchange rates and applied three models**: SARIMA**, **Prophet**, and **LSTM**. Each model offers a unique perspective on forecasting. SARIMA captures seasonal and trend patterns in time series data, making it ideal for analyzing cyclical economic behavior. Prophet, known for its flexibility, handles missing data and integrates external regressors effectively. LSTM, a deep learning model, excels in identifying long-term dependencies in sequential data. By comparing the performance and interpretability of these models, we are going to determine the most effective approach for forecasting gold prices in the futures market.

Factors influencing gold prices

Source for each factor is provided in title

Crude Oil Price: The researchers found that there's an asymmetric long-run adjustment exists between WTI and oil, and found a unidirectional relationship between gold and WTI, and that indicates the existence of such a pattern between these factors. The product we chose is not located in the same market as the main product, and the time it takes Crude Oil to respond to the Future Gold price is 10 minutes, hence, no adjustment of data is needed.

NYSE Index: The NYSE and gold prices often have an inverse relationship. When the stock market performs well, investors tend to move their money from safe-haven assets like gold to stocks, causing gold prices to drop. Conversely, during stock market downturns, investors seek safety in gold, driving its prices up. It is a major index that tracks all common stocks listed on the New York Stock Exchange (NYSE), providing a broad view of the market's performance, frankly, is not the same market as our main product. Analyzing the time frame of action of this index and how it affects Gold Futures, we can see an immediate reaction.

Natural Gas: The relationship between natural gas prices and gold futures is generally positive but indirect. When natural gas prices increase, often driven by factors like supply disruptions or economic demand, inflationary pressures rise. This typically leads to raising gold futures prices. However, this effect is not instantaneous; it

typically manifests over a short- to medium-term period, as inflation expectations and market sentiment adjust.

US Treasury Yields (Additional Link): Treasury yields signal U.S. borrowing costs and economic conditions, making them essential indicators. We included two key measures: ^TNX, the CBOE 10-Year Treasury Yield Index, which tracks 10-year U.S. Treasury bonds, a global interest rate benchmark; and ^TYX, the 30-Year Treasury Yield Index, reflecting long-term interest rate expectations. Generally, rising Treasury yields reduce gold's appeal, as bonds become more attractive with higher returns. Although yields fluctuate, research shows they are critical in predicting future gold prices.

S&P 500: The S&P 500 affects gold prices primarily through an inverse relationship during times of financial uncertainty. When the S&P 500 declines, indicating stock market instability, investors often seek the safety of gold, driving its price up. Conversely, during stable or booming stock market conditions, the demand for gold typically decreases, leading to lower gold prices. Additionally, macroeconomic factors like inflation, interest rates, and monetary policies can influence both the S&P 500 and gold prices, sometimes causing them to move in the same direction.

Ex. Rate(US/CNY): The relationship between the exchange rate (USD/CNY) and gold futures is significant and multifaceted. An increase in the USD/CNY exchange rate typically leads to a decrease in gold prices. This inverse relationship can be attributed to the fact that when the USD appreciates against the CNY, gold, which is priced in USD, becomes more expensive for investors using other currencies, thereby reducing demand and lowering prices and the inverse holds. Additionally, the USD/CNY exchange rate affects gold prices through its impact on market speculation. A stronger USD often leads to lower inflation expectations, reducing the appeal of gold as an inflation hedge. In contrast, a weaker USD can spur inflation fears, making gold a more attractive investment.

Iron: Rising iron prices often signal strong industrial demand and economic growth, leading to higher inflation expectations. This inflationary pressure can drive investors to seek gold as a safe-haven asset, thereby increasing gold futures. Moreover, the overall sentiment in the commodity markets is affected by iron prices, prompting investment shifts. When iron prices rise, indicating robust economic activity, investors may hedge against future inflation by buying gold. Conversely, falling iron prices can suggest an economic slowdown, further encouraging investment in gold for stability.

VIX : VIX, or Volatility Index, measures the market's expectation of future volatility, often referred to as the "fear gauge," reflecting investor sentiment about potential price swings in the stock market. When the VIX rises, it signals increasing investor anxiety, often due to anticipated market downturns or economic uncertainty. This heightened fear typically drives investors toward safe-haven assets like gold, leading to increased demand and, subsequently, higher gold futures prices, and the inverse is true.

Silver prices have a strong and often positive correlation with gold prices. This is due to their classification as precious metals and their similar appeal to investors, especially in times of economic uncertainty. The relationship between silver and gold often exhibits a co-movement pattern, as noted in studies that have observed cointegration between the two metals. In periods of high market demand for precious metals, silver prices increase, which can lead to a corresponding increase in gold prices as investors seek alternative stores of value.

Exchange Rate (USD/EUR) impacts gold prices primarily through its effect on the dollar's relative strength. When the dollar appreciates against the euro, it makes gold, priced in dollars, more expensive for eurozone investors, thus potentially reducing demand and lowering gold prices. Conversely, a weaker dollar typically raises the appeal of gold as an inflation hedge and alternative asset, driving up prices. This inverse relationship can be critical in periods of economic fluctuation, where shifts in the USD/EUR exchange rate can amplify movements in gold prices.

Exchange Rate (USD/CND): This research uses daily data over 10 years to investigate the relationship between USD/CND Ex. Rate and Gold Futures prices. In many time periods, the researcher found an inverse relationship, higher gold futures prices coincide with a stronger CAD relative to the USD.

Data:
For our analysis and forecasting, we focused on the closing price as the primary variable. The closing price provides a consistent and reliable measure, capturing the final market sentiment for each trading day, which is particularly relevant for identifying trends in time series data.
Source:
The data was sourced from the Yahoo Finance API, with a daily frequency covering the period from 2010-10-14 (the earliest common date across all variables) to 2024-12-05.

Preprocessing & Missing Data:
During the data review, we identified a few missing entries, which were attributed to several factors:

- Market Hours and Holidays: Some financial instruments, such as commodities like crude oil or forex, have data available outside standard stock market hours, whereas stock indices, like the NYSE Index, are unavailable during market holidays.
- Variable Update Frequencies: The frequency and timing of data updates differed across variables.

To address these issues, the following steps were taken:

1. **Linear Interpolation:** Missing data points were filled using linear interpolation. This method was chosen because the number of missing entries was minimal, making it a reasonable assumption that the interpolation would not significantly impact the analysis or model performance.
2. **Alignment:** All variables were aligned by date to ensure consistency across the dataset. This step was crucial to maintain the integrity of the time series data, as misaligned variables can introduce errors into the analysis.
3. **Scaling*:** For our deep learning models, we used the *StandardScaler* to scale the data. This was done to standardize the features, ensuring they all have a mean of 0 and a standard deviation of 1.
4. **Data Representation:** each data point has the *input* which is the data from the past 9 days and *target* which is the value of gold on the 10th day

In total, we have 3,560 data points.

Scaling* is crucial in deep learning because it prevents certain features with larger numerical ranges from dominating the training process, allowing the model to learn more efficiently and converge faster. For the model's input, we provided data from the past 9 days and tasked the network with predicting the value on the 10th day. This approach helps capture temporal dependencies and patterns in the data, enabling the model to forecast future gold futures prices based on historical trends.