Statement of Work

Gold Price Predictor

Sahiba Sachdeva (100801804)

Rationale Statement

The gold standard is a monetary system where a country's currency or paper money has a value directly linked to gold. A country that uses the gold standard sets a fixed price for gold and buys and sells gold at that price. That fixed price is used to determine the value of the currency. Import and Export of gold affects the currency of a country i.e. countries that are large importers of gold will inevitably end up having a weaker currency when the price of gold rises.

Since gold plays such an important role in deciding the value of currency of any country. It becomes a major form of investment.

During the Inflation, the relative purchasing power of gold soars while other prices dropped sharply. This is because people chose to hoard cash, and the safest place to hold cash was in gold and gold coin at the time.

Through the years, gold prices have served as a hedge against inflation and the erosion of major currencies, and thus is a fruitful investment.

In any investment it is important to know the right time when to invest the money so that it becomes profitable.

Proposed Solution

Now, gold is not just a precious heavy metal, it is a form of investment. And knowing it is right time to invest is an important factor.

The proposed project uses the machine learning algorithms to give the predictions of gold prices telling whether to invest to that specific day or not based on the data collected over the years.

Data Requirements

For the creation of predictor, we would be requiring the prices of gold over the years so that we have sufficient data for our system.

We would be acquiring the data from 'Yfinance' library. yfinance is an open source library developed by Ran Aroussi to access the financial data available on Yahoo Finance.

This library has the data from various stocks, but we would be focusing on the data of GOLD.

```
: pip install yfinance
```

Initially, it is important to install the yfinance library, then import it. The data has the daily prices of the gold.

```
import yfinance as yf
Df = yf.download('GLD', '2010-01-01', '2020-10-30', auto_adjust=True)
Df.head()
```

```
[**********************100%**********************]  1 of 1 completed
```

Date	Open	High	Low	Close	Volume
2010-01-04	109.820000	110.139999	109.309998	109.800003	16224100
2010-01-05	109.879997	110.389999	109.260002	109.699997	14213100
2010-01-06	110.709999	111.769997	110.410004	111.510002	24981900
2010-01-07	111.070000	111.290001	110.620003	110.820000	13609800
2010-01-08	111.519997	111.580002	110.260002	111.370003	15894600

The Process
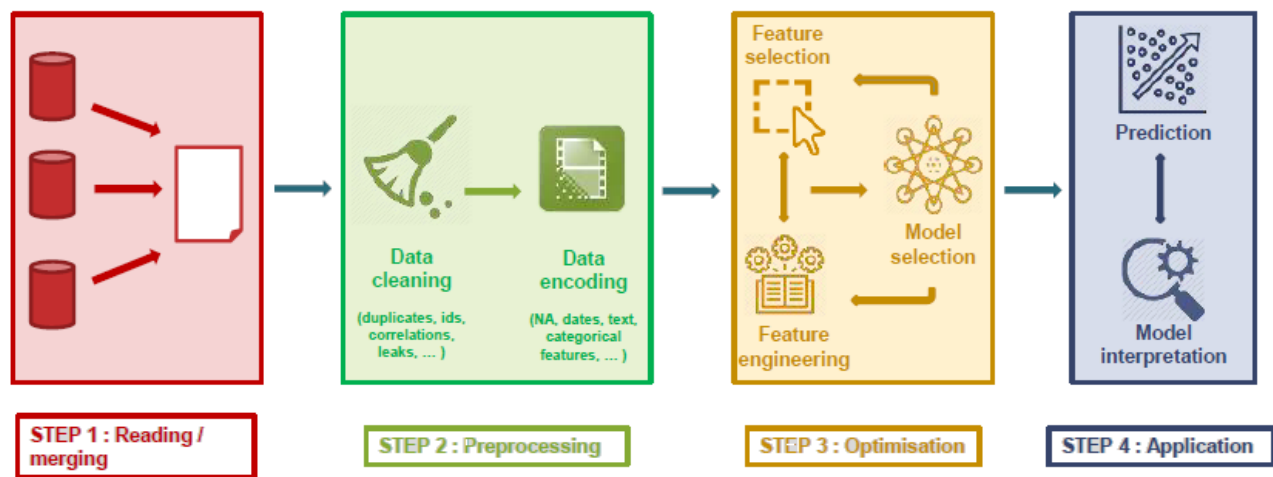
The project would be done in 4 steps with each step being a crucial step and highly dependable for the following steps. Here is a summary of each step:

1. Firstly, the data is being read the from library and is merged with other data if needed. The read data would be stored in a data frame.

2. After data is stored in the data frames, we would clean the data i.e. remove redundancy, null values etc. making sure that we have correct data for proper results of our algorithm implementation.

3. Once data is being cleaned, we would select the inputs i.e. features for our model along with selecting the best resulting machine learning algorithm for our problem statement. In this step, we would be doing feature engineering, feature selection and model selection. Before we process our data, it is very important to split the data into training and testing data, it would prevent data leakage.

4. Once we are done with the data processing, featuring engineering, the next step would be to use the data in the algorithm to get the predictions. We would be using the training data to train the algorithm and the test data would be used for testing.

Note – The parameters and approach listed above are tentative and are subjected to change as per the change in course plan and/or as per the betterment of the project. The plan and the statement of work would be updated as per the changes made.