RESEARCH ARTICLE

Gold Exchange Traded Funds (ETFs), Take Profit / Stop Loss Prediction using Machine Learning

Tamer Sh. Mazen
Assistant Professor, Dept. of Management Information Systems, Modern Academy for Computer Science and Management Technology
Cairo, Egypt

Corresponding Author: Tamer Sh. Mazen, E-mail: tamer.mazen@cs.modern-academy.edu.eg

ABSTRACT

Investment vehicles called Gold-Traded Funds (ETFs) try to follow the price of gold and are traded on stock markets. Without having to actually buy or store the metal, they provide investors with access to gold. Like stocks, gold ETF shares may be purchased and sold. Stop loss and take profit are common risk-management techniques for trading and investing that are automatically generated when the price of an asset rises above or below a certain threshold. Stop loss orders are used to minimize possible losses by closing out a trade before the price falls any lower. All earlier mentions combined; Our research showed that in order for an investor to get the maximum profit or the smallest loss, the movement of the points used to halt a loss or take gains must be adjusted. In our study, we used a series of algorithms to anticipate the market's direction and the progress in that direction. The points were adjusted automatically based on the outcomes. This study demonstrates how to establish an easy-to-use machine-learning model that aims to forecast the gold market, Stop Losses and take profit. We shall predict market trends to determine stop loss or take profit pips based on supervised machine learning models. Yahoo Finance released the dataset that was utilized to forecast market trends and (Stop Loss / Take Profit) Pips. Tree, Support Vector Classifier (SVM), random forest, Neural Network and K-Nearest Neighbour (KNN) Classifier are used with an accuracy of 83%, 56%, 91%, 99% and 92% in sequential.

KEYWORDS

Stop Loss, Take Profit, Gold Market, KNN, SVC, Decision Tree Algorithm, Neural Network, random forest

ARTICLE INFORMATION

RECEIVED: 20 June 2023
PUBLISHED: 30 June 2023
DOI: 10.32996/jbms.2023.5.4.2

1. Introduction

The term "gold market" describes the exchange of gold as a good or service. Due to its scarcity, tenacity, and aesthetic appeal, gold has been prized as a valuable metal for a very long time. Although gold is frequently used in jewellery, it also acts as a store of wealth and protection against inflation and financial instability. (Shahriar & Erkan, 2010)

The gold market is accessed in a variety of ways, such as:

- Gold is purchased and sold on the spot market for rapid delivery. Economic data, geopolitical events, supply and demand dynamics, investor attitudes, and economic indicators all have an impact on spot market pricing. (Shahriar & Erkan, 2010)
- Standardized agreements to buy or sell a certain amount of gold at a specified price and future date are known as gold futures contracts. (Shahriar & Erkan, 2010)
- Participants can make predictions about how the price of gold will change in the future on the futures market. The holder of an options contract has the option, but not the duty, to purchase or sell gold at a certain price within a specified time period. (Shahriar & Erkan, 2010)
- Exchange-Traded Funds (ETFs): Gold ETFs are investment funds that are traded on stock exchanges and are designed to follow the price of gold. They provide investors with access to gold without requiring them to keep or possess the metal themselves. Similar to stocks, gold ETF shares may be purchased and sold. (Shahriar & Erkan, 2010)
- Physical Bullion: Some investors choose to purchase actual gold bars, coins, or jewellery. They have the option of buying gold from licensed merchants, banks, or bullion dealers. Physical gold ownership allows investors direct possession of the metal but also comes with storage and security issues. (Shahriar & Erkan, 2010)
- Global economic circumstances, interest rates, inflation, currency fluctuations, central bank policies, and geopolitical tensions are just a few of the variables that have an impact on gold prices. When the economy is unclear or the markets are volatile, gold’s price may increase since it is frequently seen as a safe-haven asset.

It’s critical to remember that there are dangers associated with investing in the gold market, and prices might fluctuate. When thinking about investing in gold or any other financial instrument, it is advised to do extensive research, get expert guidance, and diversify your investment portfolio. (Hubert, 2020)

Stop loss orders are frequently used to create an exit plan and safeguard against downside risk. Investors can specify the maximum amount they are prepared to lose on a certain investment or transaction by setting a stop loss in the event that the market swings against their position; losses are minimized. When placing a stop loss order, the investor or traders selects a stop price at which they are prepared to sell the investment. If the security’s price reaches or falls below the stop price, the order is triggered, and the security is sold at the going market price. (Roberto & Tommaso, 2019)

The next sections of the paper will be as follows. In Section 2, the use of stock market indicators is illustrated. In Section 3, the survey of related works is presented. The proposed model is explained in Section 4, including the problem statement, followed by the experimental results for each algorithm and comparison-based accuracy in Section 5. Finally, the conclusion and future work are presented in Section 6.

2. Background
2.1. Technical analysis indicators
It is a technique for assessing and projecting future price movements of equities by scrutinizing previous price and volume data. Its foundation is the idea that previous pricing patterns and trends can shed light on the path that prices will take in the future. Technical analysts examine stock charts using a variety of tools and methods to spot patterns, trends, and indications. Technical analysis is limited since it doesn’t include important elements like a company’s financial health, market trends, or current happenings. Consequently, before making trading or investing choices based on technical analysis, traders and investors should undertake extensive research, engage in risk management, and take into account a variety of criteria. (Yong, Bo, Wen, & Wei, 2022)

2.1.1 Moving Average
To gain a broad picture of market direction, a moving average is employed. Forecasting long-term trends is where it performs best. Financial market professionals frequently use it for durations of 50 or 200 days over the long term and 7, 14, or 30 days over the short term. The average value of data over a predetermined amount of time is displayed via a moving average. Simple Moving Averages (SMA) and Exponential Moving Averages (EMA) are two of the many types of moving averages that are available. (Becket, 2019) (Heitkoetter, 2020) (Pring, 2000). A purchase signal is produced by the phenomenon known as an EMA crossover, which occurs when the short-term moving average crosses the long-term moving average. (Magda Fayek, 2013)

Eq. (1) represents how the SMA is calculated. It is computed by averaging the closing of \( n \) days. This average is moving since, at the conclusion of each trading day, the most recent day is added while the earliest day of the prior average is discarded.

\[
SMA_i = \frac{\sum_{j=1}^{i} \text{close}(j)}{n} \quad \text{Eq. (1)}
\]

Where,
- \( i \) is the period width,
- \( n \) is the period for calculating the average, and \( \text{close} \) is closing price for day \( i \).

The most recent data points are given more importance and weight by EMA. (Magda Fayek, 2013)

\[
EMA_i = (\text{close}_i \times \text{weight}_i) + \text{weight}_{MA} \times EMA(i-1) \quad \text{Eq. (2)}
\]

Similar to stocks, gold ETF shares may be purchased and sold. (Shahriar & Erkan, 2010)
Gold Exchange Traded Funds (ETFs), Take Profit / Stop Loss prediction using Machine Learning

weight_i = \frac{2}{n+1} ........................................................................................................ Eq. (3)

weight_i = 1 - \text{weight}_{i-1} ........................................................................................................ Eq. (4)

Where;

\(i\) is current day;
\(n\) is the period for calculating the average, \text{weight}_i\) is the weight assigned to the last day of calculation (day); and \text{weight}_{MA}\) is the weight assigned to the previous EMA. (Pring, 2000) (V.Kapoor, 2011)

2.1.2. Stop Loss and Take Profit

Stop loss is a strategic exit plan and safeguards against downside risk. It is crucial to remember that stop loss orders are not infallible, even if they might help reduce losses. In some instances, the execution price of a stop loss order may fluctuate considerably from the stop price, such as during periods of strong market volatility or trading gaps, leading to a greater loss than anticipated. Before using stop loss orders or any other investing plan, traders and investors should think about these aspects and speak with financial experts. (Vezeris, Kyrgos, & Schinas, 2018)

Take Profit is the distance between the traders’ entry price and their profit-taking, i. It shows how many pips (the smallest price change in the forex market) are involved. (Vezeris, Kyrgos, & Schinas, 2018)

2.1.3. The Average True Range (ATR)

It is a particular category of technical indicator used in the stock, currency, and commodities markets. It measures the volatility or usual price range of an asset over a specific period of time. The ATR is built using price data, typically using a moving average or exponential moving average (EMA) approach. The indicator considers the true range, which is the highest of the next three figures. The term “high minus low” refers to the difference between the highest and lowest pricing at the moment. “High minus the previous close” refers to the difference between the closing price during the prior period and the highest price during the current period. The difference between the closing price of the prior period and the lowest price of the current period is known as “low minus the previous close.” (Keyue, Yimeng, & Ying, 2023)

Note: Different moving averages, such as the simple moving average (SMA) or the exponential moving average (EMA), can be used to determine the ATR. Depending on a trader’s inclination or individual trading strategy, a different moving average method may be chosen. (Keyue, Yimeng, & Ying, 2023).

The following steps are commonly used to compute the ATR:

1. Calculate each period’s True Range (TR) using Eq. (5)

\[ TR = \begin{cases} \text{Max}[(\text{Low} - \text{High}), \\
\text{Abs}(\text{Low} - \text{Previous Close}), \\
\text{Abs}(\text{High} - \text{Previous Close})] \end{cases} \] ........................................................................................................ Eq. (5)

Where ;
Low the lowest price for gold per day.
High the highest price for gold per day
Previous Close last day close price for gold
Abs the absolute value

Select a certain time frame for the ATR calculation. Usually, there are 14 or 20 predetermined intervals in this case.

2. Determine the Average True Range (ATR) for the selected time frame :

\[ \text{ATR} = \frac{(TR_1 + TR_2 + \ldots + TR_n)}{n} \] ........................................................................................................ Eq. (6)

Where;
n is the total number of periods utilized and
\(TR_1, TR_2, \ldots, TR_n\) are the True Range values for each corresponding period. (Wong & Guppy, 2021)
2.1.4 Relative Strength Index (RSI)  
Most technical analysts and traders rely on the RSI, which, depending on the style of trading, may be a beneficial tool. It is employed to assess how well a currency has done over a specific time frame. Its foundation is the measurement of the variation between the magnitude and the rate of price change. (Brown, 2021) (Magda Fayek, 2013).

\[
\begin{align*}
RS &= \frac{\text{the total gains calculated over the last } n \text{ days}}{\text{n days are the total losses calculated over the last } n \text{ days}} \\
RSI &= 1 - \left[ \frac{100}{1 + RS} \right]
\end{align*}
\]

3. Literature Review  
Dimitrios Vezeris et al. (Vezeris, Kyrgos, & Schinas, 2018) analyze and contrast the outcomes of several techniques that have been added to a basic MACD automated trading system. It is found that basic MACD techniques are superior to quicker MACD take profit signals. Of the several ATR stop loss strategies, the sliding and variable ATR window produces the best results for a period of 12 and a multiplier of 6. For the first time, limits for Stop Loss are determined using an adaptive MACD expert adviser.

Jianwei E et al. (Jianwei E, 2019) proposed a new strategy called “ICA-GRUNN” Based on an analysis of independent component and neural network recurrent unit gate techniques. Statistically independent components (ICs) are extracted from the time sequence using VMD to deconstruct the original temporal series into a virtual multichannel mixed signal. To get the independent component prediction series (ICPs) and the gold price prediction outcome, GRUNN is applied to ICs. Comparative studies show that ICA-GRUNN surpasses other benchmark techniques and offers predictions with a high degree of accuracy.

Hong-Xia et al. (Hong-Xia, 2023) present three clever optimization algorithms that have been used to anticipate the rise and fall of gold and Bitcoin and the time needed for each technique, annualized return, and Sharpe ratio of the return curve derived using the algorithms have been compared. The quickest algorithm is the AC algorithm. The AC algorithm’s income curve has shown strong annualized income and the greatest Sharpe ratio. Under the premise that the gold and Bitcoin transaction fees changed, the control variable approach was utilized to optimize the position using their trading strategy and compare the return rates with the Sharpe ratios. It was discovered that a slight adjustment in costs had little impact on the final Sharp ratio, indicating that the optimization of positions and the trading outcomes were unaffected.

Tamer Sh. Mazen (Mazen, 2021) illustrates the goal functions Sharpe ratio and yearly profit used in this study’s innovative technique to improve the parameters of a group of technical analysis indicators based on a genetic algorithm (GA) for forex (foreign exchange).

Ambrish Mishra et al. (Mishra, Anand, & C. Debnath, 2023) proposed a model based on Bollinger bands, Fibonacci retracement indicators, and Heikin-Ashi features, as well as pattern and volume data recorded in the ontology as classes, attributes, and instances, are all included in a suggested framework for ontological risk reduction. The findings demonstrate that the proposed approach outperforms current indicators in terms of generating high returns and reducing risk.

Lauguico (Sandy Lauguico, et al., 2019) illustrate by utilizing Bollinger bands; it was suggested to create a fuzzy logic-based trading system for the stock market. The proposed algorithm is subjected to a number of trading strategies, such as the use of technical indicators to determine the strength of buy, hold, and sell signals; the use of Bollinger bands, which use stock price data to display opening and closing prices; and the classification of membership functions based on the traders’ input parameters.

Nishu Sethi (Nishu , Neha , Juhi , & Praveen , 2020) forecast market direction trends using Fibonacci retracement. By using the Golden Ratio, which is meant to provide a sound structural architecture for the stock market, they claimed that the Fibonacci retracement gives the present time frame of entry and exit locations in 70% of situations. This indicator’s theory, which predicts stock values for the near future using technical analysis, is supported by data. They claim that by applying the Fibonacci retracement formula to the two points (high and low) on the graph, it may be determined manually or automatically. In order to improve dependability, this indicator is used in conjunction with other fundamental tools to forecast both points (the entrance and exit points) in the stock market.
4. Methodology

4.1 Problem statement
From all earlier sources, in order for the investor to make the biggest profit or the lowest loss, we discovered that the movement of the points used to halt the loss or take gains must be adjusted. In our study, we used a series of algorithms to forecast the market’s direction and directional movement. The points were adjusted automatically based on the outcomes.

4.2 Problem Solving phases
As seen in the figure,1 The problem-solving phases are represented in the following points, starting with data and ending with results.

![Problem solving model diagram](image)

4.2.1 Load online data
Historical data is downloaded from Yahoo Finance until May 2023. The dataset’s features are open, closed, high, and low prices.

4.2.2 Add new features
In this phase, new calculated features are created. New features are ALT, Double Exponential Moving Average Crossovers "DEMAC", Relative Strength Index "RSI", and Moving Average Relative Strength Index "MARI". The period for DEMAC, RSI, and MARI is calculated based on the results for the optimization model that are represented by (Mazen, 2021)

4.2.3 Define target
For our machine learning model, we must now provide the target variable. We shall instead divide the market’s tendency into two groups because forecasting real prices may be exceedingly difficult. The function figures out the highest and lowest values for each bar within the defined range and assesses the state of the market, which was in class for Up-trend and other Down-trend. Upon a predicted trend, the pip value for take profit and stop loss will be determined.

Figure. 2 illustrates gold close prices for the selected period. And EMA for closing price with periods of 40, 80, and 160 days.

![Gold/USD Close Price and Moving Average](image)
Figure 3 illustrates RSI values and their borders, where a 30 border means the price is in a supported area, and a 70 border means the price is in a resistance area.

A function that accepts a set of data from an array and returns the slope of a linear regression line that fits those values is created. Then utilize this function to get the slopes of the moving averages over time, as seen in Fig. 4 (MA40, MA80, and MA160), the average price (Average), and the relative strength index (RSI). We then depict the estimated moving average slopes.

4.2.4 Splitting features and target
Currently, the data must be prepared for algorithms, which include the characteristics and targets that were separated from the data. The fact that we won’t be choosing the close, open, high, or low prices in absolute terms is crucial to keep in mind because using only these basic parameters would prevent the model from being able to predict outcomes accurately. Another way to look at it is that you would not be able to forecast the pattern if you were given a random figure for the currency’s exchange rate, which is due to the fact that the trend doesn’t always depend on a single figure at a particular time but rather on the last several days taken together. Then, we employ the characteristics we developed earlier (ATR, MV, and RSI).

4.2.4 Splitting train and validation sets
Finally, we should train the model on a subset of the data and save the remaining portion for validation in order to assess the model’s performance. In the past, we would take the entire dataset, shuffle it, and choose samples at random for training and validation. But in our situation, the qualities we developed are very consistent.
4.3 AI Algorithms
Due to the problem’s accuracy, particularly in relation to business activities that have a profit or loss, and as we know, gold’s price moves quickly. As seen in Fig. 1, a variety of models have been used to select the best one to predict the market trend. We wish to have the maximum profit or the lowest loss.

4.4 Performance Evaluation Metrics
4.4.1 Decision Tree Algorithm
As seen in the figure 5, and Table 1, a decision tree algorithm segmentation is represented with a confusion matrix that provides a thorough analysis of the results and lists the accurate and wrong predictions made for each class. By contrasting the expected outcome values with the actual values, the performance is evaluated. Where from 2314 elements. The actual downtrend elements are 1233 and 1081 for the uptrend, and the predicted downtrend elements are 1262 and 1052 for the uptrend.

Figure 5 Decision tree algorithm segmentation.

Table 1 Decision Tree Algorithm Confusion Matrix

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Down Trend</th>
<th>Up Trend</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Down Trend</td>
<td>85.5 %</td>
<td>14.6 %</td>
<td>1233</td>
</tr>
<tr>
<td>Up Trend</td>
<td>14.5 %</td>
<td>85.4 %</td>
<td>1081</td>
</tr>
<tr>
<td>Count</td>
<td>1262</td>
<td>1052</td>
<td>2314</td>
</tr>
</tbody>
</table>

4.4.2 Support Vector Machine Algorithm
As seen in Figure 6, and Table 2, a support vector machine algorithm is represented with a confusion matrix, Where from 2314 elements. The actual downtrend elements are 1233 and 1081 for an uptrend, and the predicted downtrend elements are 1268 and 1046 for an uptrend.
Figure 6. Support Vector Machin Algorithm Segmentation

Table 2. Support Vector Machin Confusion Matrix

<table>
<thead>
<tr>
<th>Actual</th>
<th>Predicted</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down Trend</td>
<td>59.8 %</td>
<td>1233</td>
</tr>
<tr>
<td>Up Trend</td>
<td>45.4 %</td>
<td></td>
</tr>
<tr>
<td>Count</td>
<td>1268</td>
<td>1046</td>
</tr>
</tbody>
</table>

4.4.3 Random Forest Algorithm

As seen in Figure 7, a random forest algorithm is represented. The algorithm is run by splitting the dataset into 2 random trees. The confusion matrix is seen in Table. 3 is presented that, from 2314 elements. The actual downtrend elements are 1233 and 1081 for an uptrend, and the predicted downtrend elements are 1438 and 876 for an uptrend.

Figure 7. Random Forest Algorithm
### 4.4.4 Neural Network Algorithm

As seen in Figure 8, the neural network algorithm is represented. The input layer neurons are ATR, MV40, MV80, MV160 and RSI. The confusion matrix is seen in Table 4 is presented that, from 2314 elements. The actual downtrend elements are 1233 and 1081 for the uptrend, and the predicted downtrend elements are 1257 and 1057 for an uptrend.

![Figure 8. Neural Network Algorithm](image)

#### Table 4. Neural Network Confusion Matrix

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Actual</strong></td>
<td></td>
</tr>
<tr>
<td>Down Trend</td>
<td>75.2 %</td>
</tr>
<tr>
<td>Up Trend</td>
<td>24.8 %</td>
</tr>
<tr>
<td>Count</td>
<td>1438</td>
</tr>
</tbody>
</table>

### 4.4.5 K-Nearest Neighbor

As seen in Figure 9, the K-Nearest Neighbor algorithm is represented. The confusion matrix is seen in Table 5 is presented that, from 2314 elements. The actual downtrend elements are 1233 and 1081 for an uptrend, and the predicted downtrend elements are 1260 and 1054 for an uptrend.

#### Table 5. K-Nearest Neighbor Confusion Matrix

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Actual</strong></td>
<td></td>
</tr>
<tr>
<td>Down Trend</td>
<td>94.0 %</td>
</tr>
<tr>
<td>Up Trend</td>
<td>6.0 %</td>
</tr>
<tr>
<td>Count</td>
<td>1257</td>
</tr>
</tbody>
</table>
Table 5. K-Nearest Neighbor Confusion Matrix

<table>
<thead>
<tr>
<th>Actual</th>
<th>Predicted</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Down Trend</td>
<td>Up Trend</td>
</tr>
<tr>
<td>Down Trend</td>
<td>90.0 %</td>
<td>9.4 %</td>
</tr>
<tr>
<td>Up Trend</td>
<td>10.0 %</td>
<td>90.6 %</td>
</tr>
<tr>
<td>Count</td>
<td>1260</td>
<td>1054</td>
</tr>
</tbody>
</table>

5. Results
Table 10. shows the accuracy for the percentage of positive identifications (Precision). The performance of the proposed classification model by using all classification thresholds is represented by a Receiver Operating Characteristic (ROC) curve in Figure 10. Finally, we found that neural networks had the best level of accuracy, with 99%.

Table 6. Machine Learning Models and Accuracy

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree</td>
<td>83%</td>
<td>85%</td>
</tr>
<tr>
<td>SVM</td>
<td>56%</td>
<td>57%</td>
</tr>
<tr>
<td>Random Forest</td>
<td>91%</td>
<td>78%</td>
</tr>
<tr>
<td>Neural Network</td>
<td>99%</td>
<td>94%</td>
</tr>
<tr>
<td>KNN</td>
<td>92%</td>
<td>90%</td>
</tr>
</tbody>
</table>
6. Conclusion
The demand for jewellery investments (including gold ETFs, bars, and coins) plays a big part in the supply side of the market. Monitoring economic metrics like GDP growth, inflation rates, and central bank policies can help determine the possible effect on gold prices. Significant gold reserves are held by central banks, especially those of big economies. Gold-specific technical analysis techniques and indicators may be used to spot possible price patterns and market trends. When making an investment choice involving gold, it is crucial to undertake extensive research, keep up with current developments on the international stage, and take into account both fundamental and technical analysis. Stop Loss and Take Profit are risk management strategies that are frequently used in trading and investing to limit losses and secure gains. They are conditional orders that are generated by investors when a trade is opened or after a position has been taken. They are useful basic tools for risk management. Orders may be executed at a different price than expected due to market factors, including price slippage and gaps in pricing. The stop loss and take profit levels should be adjusted carefully based on the trader’s risk tolerance. This research studies the market trend-based technical analysis indicators ATR, MA and RSI using some machine learning models, i.e., Tree, SVM, random forest, Neural Network, KNN and Gradient Boosting. Based on the market trend, the stop loss and take profit signals are established. The proposed model provides an adaptive trading plan based on the trend-motivation power so that the number of pips used and their positions are determined according to the updated plan. Results showed that the ANN is the best model, with an accuracy of up to 99%.
6.1 Future Work
We suggest using AI with Elliott waves and Gan analysis to predict trends, especially for long term investments.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References