On predictability of precious metals towards robust trading

Turan Erman Erkan[1,*], Adil Gürsel Karaçor[2]
Department of Industrial Engineering, Atilim University, Ankara, Turkey[1]
Grodan BV, Roermond, the Netherlands[2]
erman.erkan@atilim.edu.tr

Abstract: *Large amounts of liquidity flow into several precious metals every day. Investment decisions are mainly based on predicting the future movements of the instrument(s) in question. However, high frequency financial data are somewhat hard to model or predict as stochastic processes and many other random factors are involved. It would be valuable information for the investor if he or she knew which precious metals were quantitatively more predictable, that would also be a good basis for more robust trading decisions. The objective in this study is to build predictive models on high frequency precious metal data and compare predictabilities of different metals using only past price and volume values which should be a basis for robust trading decisions. The data used consist of various frequencies from 1-minute to 4-hour covering a period of almost 20 years for each instrument and frequency. Artificial Neural Network (ANN) and Gradient Boosted Decision Tree (XGB) methods are applied. Comparable results are achieved.*
Keywords*: PRECIOUS METALS, PREDICTABILITY, ARTIFICIAL NEURAL NETWORKS, XGBOOST*

1. Introduction

Large amounts of liquidity flow into the markets all around the world for the trade of various financial instruments such as stocks, commodities, precious metals, foreign exchange (FOREX), futures, and so on every day. Per experts and professionals, average daily turnover in precious markets alone is more than 50 billion US dollars. Investment and trading decisions, whether they tend to be long term or short term, are mainly based on predicting the future movements of the precious metal(s) in question. There has been an ongoing debate among researchers on whether financial markets are predictable or not for a long time. Some think that financial market movements are nothing but random walk, and some findings support that claim: for example, VIX futures prices were found to be unpredictable by Konstantinidi & Skiadopoulos [1]. Furthermore, some researchers even claimed that none of the conventional predictive models proposed in the literature on stock prediction seems capable of systematically predicting stock returns in long range of time horizons, and speculators do not earn significant profits in commodity and interest rate futures markets in aggregate [2, 3, 4]. On the other hand, many researchers disagree with this random walk approach. Some of them claim speculators can gain profits on commodity and currency futures [5, 6, 7, 8, 9]. Some others believe that financial instruments, commodities in particular, are predictable, at least to a certain extent [10, 11]. The debate has not been settled yet; however, it is fair to say that financial instruments are hard to model or predict, if not totally unpredictable. Obviously, it would be valuable information for the investor if he or she could know in advance which way an instrument would go. In this paper an investigation is carried out to give the investor a potential advantage by trying to answer the following question: would a certain precious metal price be likely to go up or down in the following hours? Technical (quantitative) analysis i.e. past price values of the instruments is solely used, and fundamental analysis is not in the scope of this study.

Fig. 1 *Sample OHLC price graph for platinum (XPTUSD)*

2. Data

The data used in our model consisted of 1-minute (1m), 5-minute (5m), 15-minute (15m), 30-minute (30m), 1-hour (1h), and 4-hour (4h) intraday frequencies covering the period between 2000 and 2019. High frequency data were preferred, because in lower frequencies more and more non-technical fundamental factors might start to affect price movements. Three types of precious metals were considered in this study: platinum (XPTUSD), silver (XAGUSD); and gold (XAUUSD). A sample graph of platinum (XPTUSD) is given in Figure 1.

3. Methodology

Twenty-three out of the twenty-four features used for our prediction model are related to past price movements (past high/low values, etc.), and one is related to seasonality. The output was the prediction of the actual result i.e. the maximum price value into the next eight periods (8x1-minute, 8x5-minutes, and so on).

The chosen state-of-the-art Machine Learning (ML) tools are: Extreme Gradient Boosting (XGB) method and Artificial Neural Network (ANN). Usage of ANNs is quite popular for modeling, prediction, and decision making over financial data and ANNs are regarded as an excellent tool for the purpose [12, 13, 14, 15, 16, 17, 18]. On the other hand, usage of XGB method is very rare [19].

XGB is chosen as one of the prediction methods, mainly because it has gained much popularity in the recent years, due to its overwhelming success in data science competitions [20]. Despite its reputation, XGB method has not yet achieved popularity in the literature accordingly.

ANNs are inspired from biological neural networks, and are well known for their high approximation and modelling capabilities [21, 22]. ANNs are used in a wide range of applications from job satisfaction performance evaluation [23], solving capacitated P-median problem [24] to prediction of aircraft accident occurrence [25]. ANNs are trained in order that they learn a set of input-output data that represent usually a very complex or even undefined function. With sufficient number of hidden layers and neurons, they can model any given input-output relationship [26]. All nodes (artificial neurons) are interconnected, thus form a massive parallelism, and each connection has a weight that changes as the ANN is trained, and also each node has an activation function. There are numerous activation functions, ranging from simple linear functions to various nonlinear ones. The nonlinearity of activation functions enables the ANN to learn even the most complex patterns. In this study; following a parameter optimization, an ANN model is constructed with a multi-layered feed forward network, having 2 hidden layers with 55 and 3 neurons, respectively. Hyperbolic tangent is chosen as activation function. The network is trained by back propagation gradient descent algorithm.

4. Training, Testing and Results

The data are randomized to achieve a fair distribution. 70% of the data is chosen to train the models, 20% for testing, and 10% for validation. Python software [27] is used for ANN and XGB modelling.

Both XGB and ANN models were trained until no improvement was brought about in the cross validation set. One of the main issues concerning machine learning is over-learning or over-fitting problem, in which the system memorizes a certain data set rather than learning it, because of too much training. The performance of the system can be excellent on that data set; however, it performs poorly on different data due to the lack of generalization. The opposite of this issue is the under-learning problem. Therefore, sufficient training or hyperparameter tuning is crucial for an optimum model performance. The criterion of training until improvement in the cross validation set stops is a trade off in this study.

Performances on test sets were averaged to obtain NMSE values. NMSE is mean square error divided by variance of desired output. Being a normalized value, it could easily be used for comparing different instruments of different prices and of different frequencies. Since it is an error term, values closer to zero denote better predictability.

Another statistically meaningful variable we used for predictability performance comparison is the correlation coefficient R. R is used to measure how well one variable fits on another, linear regression wise. In our case, these variables were predicted against desired. R value is calculated by the formula below:

$$R = \frac{\dfrac{\sum_i (x_i - \bar{x})(d_i - \bar{d})}{N}}{\sqrt{\dfrac{\sum_i (d_i - \bar{d})^2}{N}}\sqrt{\dfrac{\sum_i (x_i - \bar{x})^2}{N}}} \quad (1)$$

Where N = number of exemplars in the data set, x = model output, and d = desired output. The size of the mean square error (MSE) can be used to determine how well the network output fits the desired output; however, it does not necessarily reflect whether the two sets of data move in the same direction. For instance, by simply scaling the network output, we can change the MSE without changing the directionality of the data. The correlation coefficient R solves this problem. The correlation coefficient between a model output x and a desired output d is defined by formula (1). The correlation coefficient is confined to the range [-1 1]. When R = 1 there is a perfect positive linear correlation between x and d, i.e. they vary accordingly, which means that they vary by the same amount. When R = -1, there is a perfect linear negative correlation between x and d, i.e. they vary in opposite ways (when x increases, d decreases by the same amount). When R =0 there is no correlation between x and d, i.e. the variables are called uncorrelated. Intermediate values describe partial correlations. All recorded and calculated performance comparison values in terms of R and NMSE for all metals and all frequencies are given in Table 1 and Table 2 which hopefully could give hints about predictability.

Table 1: *Detailed predictability performance comparison for the precious metals on lower frequencies*

Instrument	4h		1h		30m	
	R	***NMSE***	***R***	***NMSE***	***R***	***NMSE***
XPTUSD	0.36	0.91	0.4	0.94	0.42	0.85
XAGUSD	0.41	0.84	0.44	0.81	0.46	0.79
XAUUSD	0.40	0.84	0.45	0.80	0.36	0.87

Table 2: *Detailed predictability performance comparison for the precious metals on higher frequencies*

Instrument	15m		5m		1m	
	R	***NMSE***	***R***	***NMSE***	***R***	***NMSE***
XPTUSD	0.30	0.90	0.39	0.87	0.36	0.87
XAGUSD	0.37	0.86	0.25	0.94	0.27	0.93
XAUUSD	0.36	0.87	0.38	0.85	0.27	0.93

5. Conclusion

As can be seen from Table 1 that silver (XAGUSD) has the highest overall predictability both R and NMSE wise on lower frequencies. We can say that 30m frequency has the best predictability performance again both R and NMSE wise among the lower frequencies. On the other hand, gold (XAUUSD) has the best 1h performance. As per higher frequencies which are shown in Table 2, there is no clear winner metal. Silver and gold perform similarly in higher frequencies, while platinum is considerably better in 1-minute and 5-minute frequencies. Overall metals perform better on lower frequencies. Therefore, from the investor's point of view, trading decisions based on 1-hour or 4-hour data should probably be more profitable, considering commissions and spread margins are usually more disadvantageous in higher frequencies like 5-minute or 1-minute.

Adding more metals such as copper and palladium could be considered as future work. Of course, building a robust trading strategy using predictability as a metric is also in the scope of future work.

6. References

1. Konstantinidi E., Skiadopoulos G. Are VIX futures prices predictable? An empirical. investigation. IJOF 27 543–560 (2011)
2. Bossaerts P., Hillion P. Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn? Rev. of Fin. Stud., 12(2), 405−428 (1999)
3. Goyal A., Welch I. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. Rev. of Fin. Stud. Vol. 21 Issue 4, pp. 1455-1508 (2008)
4. Hartzmark M. L. Returns to individual traders of futures: Aggregate results. JPE, 95, 1292–1306 (1987)
5. Kho B. C. Time-varying Risk Premia, Volatility, and Technical Trading Rule Profits: Evidence from Foreign Currency Futures Markets. JFE, 41, 249–290 (1996)
6. Strozzi F., Zaldivar J.M. Non-linear Forecasting in High-frequency Financial Time Series. Physica A, 353 (2005) 463–479 (2005)
7. Taylor S. J. Rewards Available to Currency Futures Speculators: Compensation for Risk or Evidence of Inefficient Pricing? Economic Record, 68(Supplement), 105–116 (1992)
8. Wang C. Futures Trading Activity and Predictable Foreign Exchange Market Movements. J. of Bank. & Fin. 28, 1023–1041 (2004)
9. Yoo J., Maddala G. S. Risk Premia and Price Volatility in Futures Markets. JFM, 11, 165–177 (1991)
10. Campbell J. Y., Thompson S. Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average? Rev. of Fin. Stu. Vol. 21, Issue 4, pp. 1509-1531 (2008)
11. Zunino L., Tabak B.M., Serinaldi F., Zanin M., Perez D.G., Rosso O.A. Commodity Predictability Analysis with a Permutation Information Theory Approach. Physica A, 390 (2011) 876–890 (2010)
12. Cheng W., McClain B.W., Kelly C. Artificial Neural Networks Make Their Mark as a Powerful Tool for Investors. Rev. of Bus. 4 – 9 (1997)
13. Dash R., Dash P.K., Bisoi R. A self-adaptive differential harmony search based optimized extreme learning machine for financial time series prediction. SAEC Vol. 19 pp 25-42 (2014)
14. Dutta S., Shekhar S. Bond-rating: a Non-conservative Application of Neural Networks. Proceedings of the IEEE Int. Conf. on NN 2 pp 443– 450 (1988)

15. Lam M. Neural Network Techniques for Financial Performance Prediction: Integrating Fundamental and Technical Analysis. Decision Support Systems, 37 567– 581 (2004)
16. Wedding D.K., Cios K.J. Time series forecasting by combining RBF networks, certainty factors, and the Box-Jenkins model. Neurocomputing Vol. 10, Issue 2, pp 149-168 (1996)
17. Xi L., Muzhou H., Lee M.H., Li J., Wei D., Hai H., Wu Y. A new constructive neural network method for noise processing and its application on stock market prediction. App. S. Comp. Vol. 15, pp 57-66 (2014)
18. Yu L., Lai K.K., Wang S. Multistage RBF neural network ensemble learning for exchange rates forecasting. Neurocomputing Vol. 71, Issues 16-18, pp 3295-3302 (2008)
19. Zięba M., Tomczak S. K., Tomczak J. M. Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction, Exp. Sys. w. Apps., Volume 58, pp 93-101 (2016)
20. Chen, T., & Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22Nd ACM SIGKDD (pp. 785-794). ACM (2016)
21. Karacor A. G., Sivri N., Ucan O.N., Maximum Stream Temperature Estimation of Degirmendere River Using Artificial Neural Network, JSIR, Vol. 66 No. 5 pp. 363-366 (2007)
22. Karacor G., Denizhan Y. Advantages of Hierarchical Organisation in Neural Networks, IJCAS Vol. 16 pp 48-60 (2004)
23. Aktepe A., Ersoz S. A Quantitative Performance Evaluation Model Based on a Job Satisfaction - Performance Matrix and Application in a Manufacturing Company. IJIE, 19(6), 264 – 277 (2012)
24. Shamsipoor H., Sandidzadeh M.A., Yaghini M. Solving Capacitated P - Median Problem by a New Structure of Neural Network. IJIE, 19(8), 305 – 319 (2012)
25. Yeoum S.J., Lee Y.H. A Study on Prediction Modeling of Korea Millitary Aircraft Accident Occurrence. IJIE, 20(9-10), 562 – 573 (2013)
26. Hornik K., Stinchcombe M., White H. Multilayer Feedforward Networks are Universal Approximators, NN, Vol. 2, pp. 359-366 (1989)
27. Rossum G. Python tutorial, Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI) (1995).