



USING MACHINE LEARNING FOR GOLD PRICE PREDICTION

P. JAGADEESH, MCA, DCA, DVR & Dr. Hima Shekar MIC College of Technology, A.P., India.

K. MAHANTHI, Assistant Professor, Dept. Of AI & IT, DVR & Dr. Hima Shekar MIC college of Technology,
A.P., India.

Abstract— This thesis investigates properties of the gold price. Two different aspects have been analysed using two different methods. The concept of fractional Brownian motion has been utilized in the search for evidence of long memory in the gold price. Further, the machine learning techniques Gradient Boosting Machine and XGBoost are used in the investigation of the relative importance of financial and economic variables in a prediction of the gold price. Both aspects are studied across multiple time periods in order to examine the potential change. We find evidence of long memory in the gold price, however not in all examined time periods from 1979 until now. The first ten years from 1979, as well as the ten years from the start of the 2008 financial crisis to today, appear to have the property of persistence, while the years in between display no evidence of long memory.

Introduction

Gold is a commodity being known to mankind for ages. The scarcity, visual appearance and physical durability have made gold valuable. Its roles range from jewelry, money, industrial usage and being an investment asset. The fact that the gold price extends beyond indicating the value of a certain amount of a commodity, make sit an interesting entity to research.

This thesis aims to further look into some of the gold price' properties, predominantly in the context of gold as an investment asset. Even though gold has been function in gas an investment asset for centuries, we are mostly interested in the last decades, following the fall of the Bretton-Woods system in the 1970s. This was the first occasion the dollar was completely independent from either gold or silver since 1792, and marks a significant change in the market for gold (Hillier, Draper and Faff, 2006). This is when the gold's role as an investment asset truly emerges, and the monetary role diminishes in the world's largest economies ,at least formally.

In the investigation of the properties of the gold price we will look at two aspects. Firstly, we will look for evidence of long memory. Secondly, we will look at gold's relationship to financial and economic factors through the creation of a predictive model.



In the search for evidence of long memory, we have applied the concept of fractional Brownian motion. This concept differs from the ordinary Brownian motion, among other things, by its characteristic Hurst exponent. By estimating this exponent, we will be able to get an indication of whether the process is persistent. The phenomenon of long memory play significant role in many fields, with implications in regard to forecasting skills and trends (Graves et al. 2017), and in this thesis we are looking for evidence that the gold price is an entity influenced by this phenomenon.

With the possible presence and implications of long memory, the further analysis of the gold price properties will be conducted through a predictive model. The utilization of supervised machine learning methods, namely Gradient Boosting Machine (GBM) and XGBoost, enables the opportunity to display the relative importance of the input variables. The input variables

Area selection of economic and financial factors and indicators, chosen on background of the disapproved relevance in literature, or simply because they appear interesting based on economic intuition. Our aim is not to examine the causal relationships but rather the relative importance in a prediction.

LITERATURE REVIEW

In order to achieve the most accurate variable importance plots, the fundamental prediction model should be as accurate as possible.

Implementing the GBM and XGBoost models are in itself not too challenging, but applying them with precision is more difficult. As described in section 2.1.2, the models need inputs



such as learning rate and number of iterations along with the input data. The process of adjusting these parameters are called parameter tuning. The tuning of model parameters is challenging, and is as many other concepts within machine learning about the bias-variance trade-off (Chen et al. 2015). A biased model is oversimplified, where the difference between the average prediction of our model and the actual value we are trying to predict is high. Variance is the variability of model prediction for a given value, which indicates the spread of our data (Singh, 2018). A model with high variance is performing well on training data, ultrahigh irroration test data.

After subdividing the dataset into three different time periods, each subset were further split in two. One part will serve as a train set, and the other as a test set. The train set will be used to fit the models, and the test set is for assessing the final prediction. The split was set to 0.75 and 0.25 using random sampling. Random sampling was used foremost to prevent the possible biases that can arise when using ordered data (Mallick, 2017). Furthermore, our aim is to compare the periods as a whole, and pooling the observations does not obstruct this.

No preprocessing such as scaling or normalization is conducted on the data set. This is due to the fact that XGBoost and

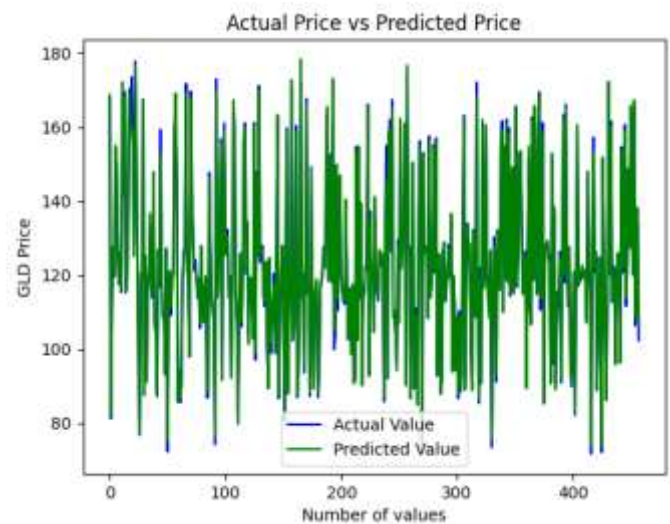


GBM with tree based learning is used in this case and is not sensitive to monotonic transformations. The models pick adequate splitting points in the variables to split at each node, where defining a split on one scale has a corresponding split on a transformed scale (Chen, 2014).

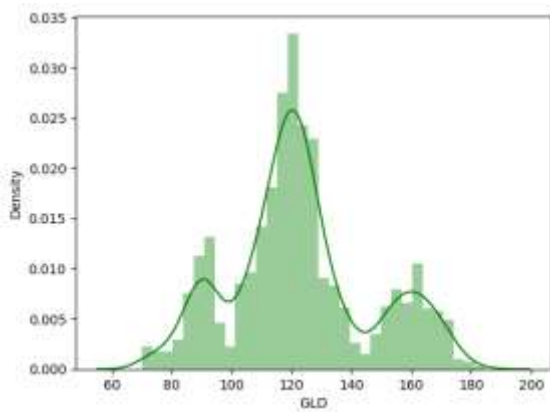
The following paragraphs describe the process of tuning the model parameters to achieve the most precise prediction models possible, in order to obtain an accurate overview of variable importance.

To overcome this problem we have developed a model which will be highly effective. Machine learning Algorithms are used because they provide us with continuous value as an output and not a categorized value. Because of which it will be possible to predict the actual price of a gold rather than the price range of a gold. User Interface has also been developed which acquires input from any user and displays the Price of a gold according to user's inputs.

SAMPLE SCREENSHOTS



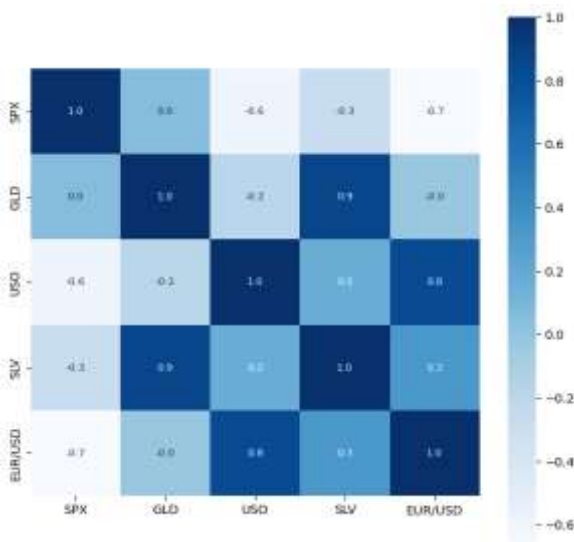
PROPOSED WORK



in a prediction of the gold price. Both parts of the analysis have been conducted on different periods in order to be able to compare the results, and assess whether they have changed over time.

REFERENCES

Aarshay, J. (2016). Analytics Vidhya. *Complete Guide to Parameter Tuning in XGBoost (with codes in Python)*. Downloaded November 1st from <https://www.analyticsvidhya.com/blog/2016/03/complete-guide-parameter-tuning-xgboost-with-codes-python/>



CONCLUSION

In this paper, we propose a investigate the properties of the gold price, and whether these have changed over time. The analysis of the properties has been twofold. Firstly, we have looked for evidence of long memory in the gold price. Secondly, we have analysed the relative importance of different financial and macroeconomic variables

Ali, S. H. (2006). Gold mining and the golden rule: a challenge for producers and consumers in developing countries. *Journal of CleanerProduction*, 14(3-4),455-462.



Aref, S. (1998). *Hurst Phenomenon and Fractal Dimension in Long-Term Yield Data*. Conference Statistics in Agriculture. Downloaded September 14th from <https://doi.org/10.4148/2475-7772.1275>

Baker, R. (2018). The Mystery of Gold and the Chinese Yuan. *Kitco.com*. Downloaded November 12th from <https://www.kitco.com/commentaries/2018-07-31/The-Mystery-of-Gold-and-the-Chinese-Yuan.html>

Batten et al. (2005). *Return Anomalies on the Nikkei: Are They Statistical Illusions?* Chaos So lit on sand Fractals 23 (4). Downloaded September 17th

from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=396680

Batten, J. A., Ciner, C., & Lucey, B. M. (2014). On the economic determinants of the gold–inflation relation. *Resources Policy*, 41, 101-108.

Baur, D. G.; Lucey, B. M. (2010). *Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold*. Financial Review, Volume 45, Issue 2.

Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking & Finance*, 34(8), 1886-1898.



Industrial Engineering Journal

ISSN: 0970-2555

Volume : 52, Issue 7, July : 2023

Beckers, S & Soenen,

L. (1984). Gold: more
attractive to non-US
than to US investors?

*Journal of Business
Finance and*

Accounting, 11, pp.

107-112