CS230

Daily Gold price forecasting using LSTM cells and Attention Mechanism CS230 - Sping 2021

Kalliopi Anastasia (Lilian) Kourti
Institute for Computational and Mathematical Engineering
Stanford University
lkourti@stanford.edu

Abstract

Commodities are of significant interest, as their prices are usually dependent on market supply and demand, and also on investor behavior and other complex economic factors. That is particularly true for Gold, which is a multifaceted asset. Gold is a commodity, a precious metal, and it is also perceived to have future purchasing power although it is not a currency. Having an accurate way to predict daily Gold prices can yield significant profits both for speculative traders and investors. Hence, we consider RNNs for this task because of their ability of finding hidden patterns in timeseries of stock prices.

1 Introduction

This project will explore daily Gold price forecasting using RNN (LSTM) networks with Attention Mechanism, which receive as input multiple daily timeseries for Gold, other related commodities and important economic factors. Also, feature selection will be carried out to find which of these features are indeed useful in this specific forecasting task.

2 Related work

A brief literature review was useful and inspiring to figure out what would be interesting to explore in this project. Feed Forward NNs and Ensemble Methods are commonly used for Gold price forecasting using multiple features [1],[2],[3]. In [4] and [5], more complex models are presented that combine CNN and LSTM components, but they only consider Gold spot price sequences as feature. Additionally, although Attention Mechanism is becoming more and more prevailing in the literature, related work on Gold price forecasting is still very limited.

3 Approach/Method

Hence, it would be interesting to experiment with Deep NN models with LSTM cells, but with multiple timeseries as input. The reasoning behind this is that RNNs, and in particular LSTM cells, are the most appropriate to capture the sequential nature of price timeseries. Thus, it sounds promising to explore if a good performance can be achieved with LSTM layers (but not necessarily CNN layers) provided that we feed the network, not only with gold prices, but also with multiple other features (see features listed in the Dataset and Features section).

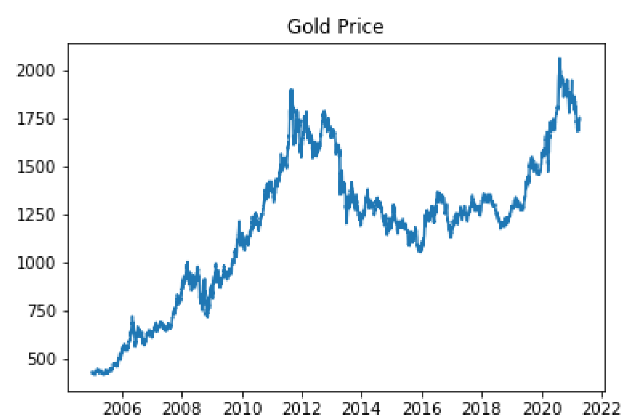
Additionally, Attention Mechanism will be added to the network as a mechanism of distribution of probability weight in order to enhance the extraction of the temporal and spatial features. Attention Mechanism, that is an evolution of the Encoder-Decoder Model for RNNs, was developed in order to avoid forgetting of the earlier parts of the sequence. The main idea is to allow the decoder to selectively access encoder information during decoding. This is achieved by building a different context vector for every time step of the decoder, calculating it in function of the previous hidden state and of all the hidden states of the encoder, assigning them trainable weights.

4 Dataset and Features

Since Gold often moves higher when economic conditions worsen, it is viewed as an efficient tool for diversifying a portfolio. Also, Gold is often used to hedge inflation because, unlike paper money, its supply doesn't change much year to year. Using the above, as well as the findings in [3], it is reasonable to conclude that:

- Gold Spot US Dollar (XAU/USD),
- Silver Spot US Dollar (XAG/USD),
- Inflation (CPIAUCNS),
- US Dollar Index (DXY), and
- US Treasury Bill rate (OMRXTBILL30)

could be meaningful predictors in forecasting Gold prices. The daily historical data for the above features can be pulled from `https://fred.stlouisfed.org/` or `https://uk.investing.com/` for the trading days during the period 01/04/2005 – 04/08/2021. The figures of each covariate can be found in the Jupyter Notebook attached to the Gradescope submission.



5 Preprocessing

The data cannot be readily fed into the model. First, the series are converted into supervised learning format based on the desired forecast horizon. A forecast horizon of 10 days is used. This means that the 9 last time steps are used to predict the value on the 10th time step. More precisely, all five timeseries are shifted by 1,2,. . . ,9 timesteps to construct the features. The result is that a sample is of the following form, where X is the input and y is the output:

X											y
Gold_price(t-9)	Silver_price(t-9)	US_dollar_index(t-9)	Tbill_rate(t-9)	Inflation(t-9)	...	Gold_price(t-1)	Silver_price(t-1)	US_dollar_index(t-1)	Tbill_rate(t-1)	Inflation(t-1)	Gold_price(t)

Subsequently, the dataset is split into training/validation/test set (60%-20%-20%) by date, and finally the inputs (X) are reshaped into the 3D format expected by LSTMs. It is noted here that because neural networks are sensitive to the diversity of input data, each feature is scaled in the range 0 to 1. Re-scaling of the output predictions is thus necessary at the end.

6 Models

The models will be developed using Keras and Tensorflow, and for the Attention Mechanism the 'attention' [1] Python package will be used. The architecture upon which the experiments will be made is a Sequential Network composed of stacked LSTM layers followed by one or more Fully Connected (i.e. FC/Dense) layers, the last of which will have a linear activation function that will yield the regression output. Additionally, variations will be made to include Attention layer before the FC layers.

7 Hyperparameter Tuning

The performance of the proposed model is highly dependent on the choice of hyperparameters. More precisely, the following hyperparameters are considered: number of LSTM layers, number of FC layers, number of neurons per layer, activation functions, batch size and number of epochs. For these hyperparameters, the following values were considered during the tuning process:

- Number of LSTM layers: 1, 2, 3
- Number of FC layers: 1, 2, 3
- Number of cells in LSTM layers: 32, 64, 128, 256
- Number of cells in Attention layer: 32, 64, 128
- Number of neurons in FC layers: 32, 64, 128
- Activation functions: tanh, ReLU
- Batch size: 256, 512, 1024
- Number of epochs: 50, 100, 150, 200

Subsequently, we present the most promising configurations. These models were picked as the best in terms of validation error. It is noted here that the Adam optimizer was used, and since the task at hand is a regression task, optimizing for minimum MSE $\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ was the most appropriate choice. To make clear the notation, it is noted that $LSTM_1(128)$ denotes that this is the first LSTM layer and it contains 128 cells. Additionally, FC_{out} is the output layer, which always has a linear activation and contains only one neuron which outputs the prediction.

- **Model 1 (validation loss 2.4926e-04)**:
  $LSTM_1$(128) - FC_1(64, ReLu) - FC_{out}
  with batch size 1024 and trained for 150 epochs
- **Model 2 (validation loss 1.9053e-04)**:
  $LSTM_1$(128) - $LSTM_2$(64) - FC_1(64, ReLu) - FC_{out}
  with batch size 512 and trained for 150 epochs
- **Model 3 (validation loss 1.3101e-04)**:
  $LSTM_1$(128) - $Attention_1$(64) - FC_1(32, tanh) - FC_{out}
  with batch size 1024 and trained for 150 epochs

8 Feature Selection

Although the timeseries that are chosen as features are correlated with the daily Gold prices, all combinations of them are explored to make sure we use the combination of features which leads to maximum performance. More precisely, the most promising model, Model 3, was trained separately on the following features combinations: (Gold, Silver, Inflation, US Dollar, US T-bill), (Gold, Silver, Inflation, US Dollar), (Gold, Silver, Inflation, US T-bill), (Gold, Silver, US Dollar, US T-bill), (Gold, Inflation, US Dollar, US T-bill), (Gold, Silver, Inflation), (Gold, Silver, US Dollar), (Gold, Silver, US

[1] https://github.com/philipperemy/keras-attention-mechanism

T-Bill), (Gold, Inflation, US Dollar), (Gold, Inflation, US T-bill), (Gold, US Dollar, US T-bill), (Gold, Silver), (Gold, Inflation), (Gold, US Dollar), (Gold, US T-bill), (Gold). From the above combinations, it turns out that Gold prices alone or in combination with Silver prices are more suitable to make accurate predictions for the daily Gold prices. More precisely, for these two cases the validation loss further drops to **2.4119e-05** and **4.2305e-05**, which constitutes a non-negligible improvement compared to the initial validation loss observed for Model 3 when trained on all five timeseries.

9 Performance and Evaluation

There are many publications on gold price forecasting that could be used to assess the model's performance. [5] is one of the most recent and achieves great performance in terms of RMSE, thus it will be used as a benchmark for this work. More precisely, our work will be compared to an CNN-LSTM network trained on gold price data for the same horizon. The models developed in our project will be evaluated (i) on a test set that corresponds to 20% of the current dataset, and (ii) on a subset of this test set that exactly matches the test set used by Livieris et al in [5]. This evaluation is presented in Table 1. [5] reports an RMSE equal to 0.0100 for their most successful model and for forecasting horizon equal to 9 (same as used in this project). Comparing this to the last column of Table 1 we can safely say that Model 3 outperforms the benchmark.

To illustrate the great performance of the developed models, Figures 1-3 are also presented and show the true and the predicted gold prices in the whole test set for Model 3 when trained on (1) all features, (2) Gold and Silver prices, and (3) only Gold prices, respectively.

	RMSE on **own** (large) test set	RMSE on **Livieris et al** test set
Model 1 (all features)	0.03835	**0.01265**
Model 2 (all features)	0.02780	**0.01034**
Model 3 (all features)	0.02058	**0.00952**
Model 3 (Gold and Silver)	0.01067	**0.00616**
Model 3 (Gold only)	0.00779	**0.00396**

Table 1: RMSE reported on model output (i.e. normalized prediction)

10 Conclusion and Future Work

In this work, RNNs with LSTM cells and Attention Mechanism were used to predict daily Gold prices. Three model architectures were explored, and different combinations of covariates were tried. Based on the results presented in Sections 8 and 9, the following are the main conclusions that can be drawn:

- LSTM architectures are successful in providing good forecasts for daily gold prices. The addition of Attention Mechanism was crucial in the improvement of the model performance.
- The chosen covariates that are correlated with the Gold price timeseries did not enhance the model's predictive capacity. On the contrary, it seems that they instilled more noise than useful signal. The only exception was the Silver price timeseries.
- The proposed architecture and configuration beat the benchmark, and achieved a good performance on an even larger test set than the one used in [5].

Future work could include forecasting using look-back and look-ahead windows of various lengths to see how the performance is affected when the model is offered more information from the past, but also it is requested to predict further into the future. More in depth research on correlated covariates would also be interesting to understand if Gold prices are indeed sufficient in forecasting Gold prices or if there exists some other covariate that could be used as feature and offer richer information.

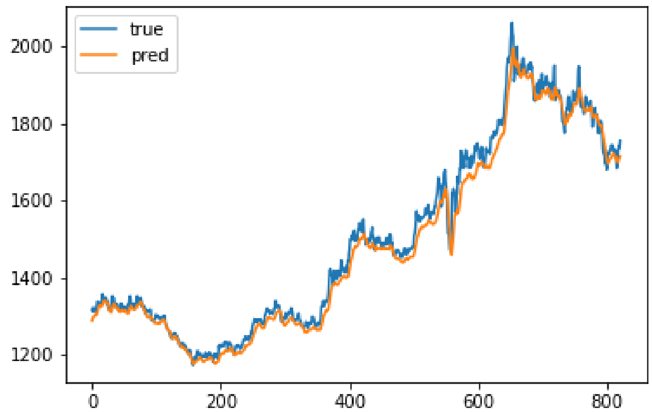


Figure 1: Model 3 evaluated on test set (all features)

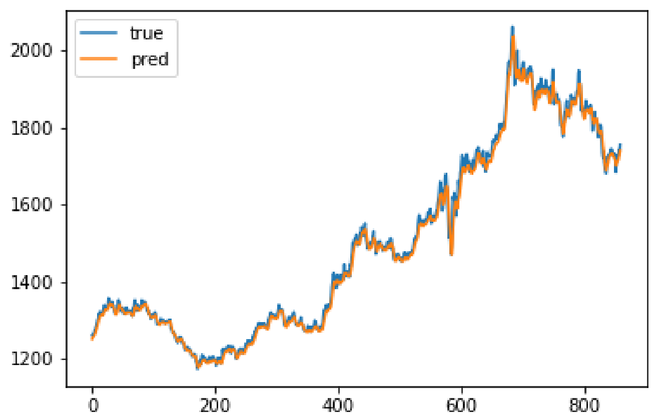


Figure 2: Model 3 evaluated on test set (Gold and Silver prices as features)

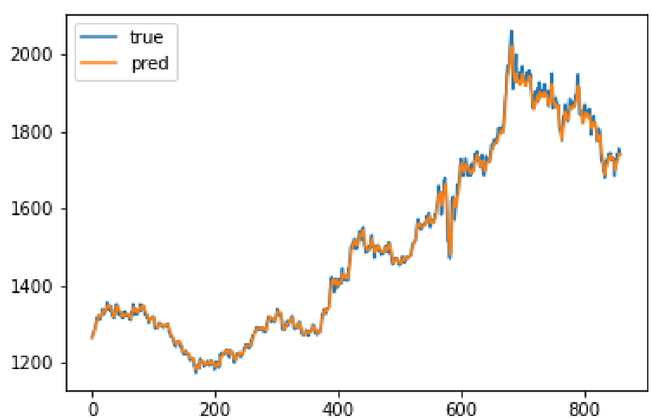


Figure 3: Model 3 evaluated on test set (Gold prices as features)

References

[1] Pandey, A. C., Misra, S., Saxena, M. (2019, August). Gold and diamond price prediction using enhanced ensemble learning. In 2019 Twelfth International Conference on Contemporary Computing (IC3) (pp. 1-4). IEEE.

[2] Mahato, P. K., Attar, V. (2014, August). Prediction of gold and silver stock price using ensemble models. In 2014 International Conference on Advances in Engineering Technology Research (ICAETR-2014) (pp. 1-4). IEEE.

[3] Ur Sami, I., Junejo, K. N. (2017). Predicting future gold rates using machine learning approach. International Journal of Advanced Computer Science and Applications, 8(12), 92-99.

[4] He, Z., Zhou, J., Dai, H. N., Wang, H. (2019, August). Gold price forecast based on LSTM-CNN model. In 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech) (pp. 1046-1053). IEEE.

[5] Livieris, I. E., Pintelas, E., Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. Neural computing and applications, 32(23), 17351-17360.