Forecasting Gold Prices Using Temporal Convolutional Networks

Justin Fajou[1][0000−0003−0597−1941] and Andrew McCarren[2][0000−0002−7297−0984]

[1] Dublin City University, Dublin, Ireland justin.fajou2@mail.dcu.ie
[2] Insight Centre for Data Analytics, Dublin City University
andrew.mccarren@dcu.ie

Abstract. Accurate prediction of the financial markets can provide many benefits, of which underlying economic stability is probably the most important. This area has understandably attracted a significant amount of interest from the research community, and has inspired a diverse range of approaches with varying degrees of success. Gold is a particular commodity which has attracted considerable attention since it was first smelted for ornaments and jewellery by the Egyptians in 3600BC. In uncertain economic times it is regularly used as a safe-haven commodity, and is why there is considerable attention given to enhancing the accuracy of gold prices prediction methods.
Previous attempts at gold price prediction have used a variety of econometric and machine learning techniques. In particular Long Short-Term Networks (LSTMs) and more recently an ensemble of Convolutional Neural Networks (CNNs) and LSTMs have been found to have had considerable level of success in time series prediction. In this research we have conducted a comparative analysis between ARIMA, CNN, LSTM and CNN-LSTM and a recently introduced structure known as Temporal Convolutional Networks (TCNs) on gold price data spanning 20 years. The results show how TCNs produced a RMSE of 15.26 and outperformed both CNN-LSTM and LSTM with RMSE scores of 23.53 and 27.39 respectively.

Keywords: TCN · Time Series · Commodities

1 Introduction

Gold was first traded as a futures contract on the Chicago Mercantile Exchange (CME), following the collapse of the Bretton Woods agreement in 1971 [27]. This event decoupled the price of gold from money, which had required central banks to purchase gold to support the value of their currencies. Subsequently, the global economy gradually moved towards a US Dollar standard and away from the gold standard. Nowadays, the US dollar is still regarded as the worlds reserve currency and the benchmark price of gold is quoted in USD. In addition to trading, the most common use for gold today is in the manufacture of jewellery, which accounts for 80% of annual demand [6]. Due to its natural properties, gold is also used in many industrial application such as mobile phones, computer equipment and even in the production of specialist glazing to reduce heating and ventilation costs in commercial buildings [6].

Throughout history, gold has always been seen as a source of value, especially in times of uncertainty and is considered to be a secure asset to store wealth. Thus gold is seen as a safe-haven investment and tends to rise when there is increased uncertainty in the world economy and drops when economic activity increases [3]. Times of uncertainty have included war, recession or even the current Covid-19 pandemic, which has seen the price of gold rise significantly [38]. In addition, gold has historically been used as a hedge against inflation, however according to [19] the link is more apparent in bear markets so the correlation is not consistent and dependent upon specific market scenarios.

The gold market is essential to a properly functioning economy, and because of its influence and importance to the global economy, the prediction of gold has received extensive attention by researchers, in order to understand the unique factors that influence price changes. For example approaches ranging from traditional econometric models such as ARIMA (Auto-Regressive Integrated Moving Average) [13] and GARCH (Generalized Auto-Regressive Conditional Heteroskedasticity) [9] models, to machine learning approaches such as Support Vector Machines (SVMs) [20] and Random Forests [25] have all been used to predict gold in the past.

More recently, deep learning architectures have shown success when forecasting time-series. Deep neural networks can be defined as having an input layer, an output layer and multiple hidden layers that can extract granular features from the input data. Artificial Neural Networks (ANNs) have been used to predict gold prices as in [30, 36]. Long Short-Term Networks (LSTMs), which are a type of Recurrent Neural Network (RNN), have shown to be very effective at sequence modelling [34]. They have also been used in conjunction with Convolutional Neural Networks (CNNs) to predict gold prices as in [15, 26, 37].

1.1 Contribution

In this paper a comparative analysis between a number of univariate prediction techniques and a relatively new deep learning approach known as Temporal Convolutional Networks (TCNs) was undertaken to predict the price of gold. This is the first examination of the predictive power of TCNs on gold prices [2]. This research will show that TCN models can be used as a viable alternative to current state of the art deep learning models when predicting gold price.

1.2 Structure of Paper

The remainder of this paper is organised as follows: section 2 provides a literature review of the current gold price prediction strategies and the current application areas for TCNs; section 3 gives a detailed description of the data and methods used in this study to forecast gold prices; section 4 gives an overview and comparison of the performance for all the models that were implemented; and finally in section 5 we summarise our outcomes and outline potential future work.

2 Related Research

In this section we initially examine the use of machine learning in stock market forecasting and then focus on the techniques used specifically in gold price prediction. Finally, we explore the the latest research on TCNs, which provide the basis for this paper.

2.1 Stock Prices

A "Feature-Fusion LSTM-CNN" was used to predict stock prices in [22]. They constructed a number of images, which were the fusion of the stock price and volume images. These images were then fed to a CNN model, which incorporated residual connections to allow for a deeper network design [14]. A separate time series of the stock and volume is then fed to a LSTM. The features were then concatenated and inputted to fully connected layers. The predictions from the LSTM-CNN outperformed both the individual CNN and LSTM models. In this research the authors collected minute by minute data points from a S&P500 Exchange-Traded Fund (ETF), which generated nearly 100k data points. Evaluation of their results was completed using the root mean squared error (RMSE), the root mean absolute error (RMAE) and the mean absolute percentage error (MAPE). Whilst they claim this approach produced very positive results, the resulting model is quite complex and the data preparation required for the chart generation would take a considerable amount of time.

A purely convolutional approach was used by Hoseinzade and Haratizadeh, who presented a method they refer to as CNNPred [17]. They attempt to classify the next days trading as "Up" or "Down" for the S&P500, NASDAQ, DJI, NYSE and RUSSELL stock market indices. The first model, called 2D-CNNPred, used two dimensional inputs that feed images to the CNN containing 60 days time lags and 82 explanatory variables. The second approach known as 3D-CNNPred, adds an additional dimension of features from the five financial indexes that are being predicted. Both models use multiple convolution and pooling layers and then flatten the data, which is fed to a fully-connected layer to produce the final output. They used F-measure to evaluate their results with both models outperforming the baselines, which included CNNs and ANNs. While they achieved positive results using a convolutional approach and their architecture didn't include LSTMs as part of their design.

2.2 Gold Prices

A whale optimization algorithm (WOA) was proposed to forecast gold prices in [1]. The WOA is a meta-heuristic designed to optimize the search space, and is used in conjunction with a multilayer perceptron neural network. There were 360 monthly gold prices used for the model, with an additional 10 predictor variables such as commodity prices, currency prices and inflation. They used mean absolute error (MAE) and RMSE to evaluate their findings. They found 85% increase in performance over baseline models which included ARIMA and other meta-heuristic combinations with neural networks. It should be noted the results are dependent on a small set of monthly gold prices.

Jianwei et al. [21] introduced a novel technique known as ICA-GRUNN (Independent Component Analysis - Gated Recurrent Neural Network) to predict the price of gold. Their method extracts features using ICA and passes these features to a GRUNN to forecast gold prices. They used gold prices from Jan 1979 to Dec 2017 to train and test the models. Evaluation of the models were performed using mean absolute deviation (MAD), MAPE and RMSE. ICA-GRUNN model outperformed the baseline models (ARIMA, LSTM and GRU (Gated Recurrent Networks)), indicating this technique warrants further research.

The goal of Vidal and Kristjanpoller was to predict the volatility of gold [37]. They proposed using a CNN-LSTM hybrid model. One significant difference to other studies is the use of a VGG16 network. VGG16 is a CNN that was pretrained on the ImageNet dataset [8]. In this study, it receives a time-series that has been transformed to a RGB image. The LSTM model used log-transforms of the price as input. Both feature spaces are concatenated and fed to a fully connected layer to predict the volatility. They used gold spot prices from the London Bullion Market Association (LBMA) from April 1968 to October 2017 and used MAE for their evaluation. Comparisons with baseline models showed an 18% improvement on LSTM models.

Similarly, Livierieris et al. proposed a CNN-LSTM method for the prediction of gold prices [26]. However, instead of concatenating the feature space, the model was generated sequentially and fed to a fully connected layer. They used gold prices from Jan 2014 to April 2018, sourced from https://finance.yahoo.com. In addition to predicting the gold price, they also produced a classification of whether the price would increase or decrease the next day. MAE and RMSE were used for evaluation and they successfully beat the performance of their LSTM base models for both gold price prediction and next day classification. However, the classification results don't appear to be much better than chance, with the best accuracy score being 55.53%, but that could be related to the small dataset.

2.3 Temporal Convolutional Networks

The development of Temporal Convolutional Networks was significantly influenced by the onset of WaveNet's [31], which proposed the use of Dilated Causal Convolutions within Residual Connections. TCNs were first used in the segmentation and detection of video and were found to be successful compared to CNNs and LSTMs on multiple datasets [24].

TCNs were first used on time series data by Bai et al. [2]. They proposed a network with three key pillars: Causal Convolutions, Dilated Convolutions and Residual Connections, and combined these three concepts, along with the more typical neural network components of Dropout, Batch Normalization and the Rectified Linear Units (ReLU) activation function to define the TCN network. In this study a combination of accuracy, loss and perplexity were used to evaluate the performance of the TCN on 11 sequencing tasks. The TCN outperformed LSTM, GRU and RNN models in 10 out of 11 of those tasks.

As this concept is relatively new, there has not been an extensive volume of research comparing the results of TCNs to LSTMs, when predicting stock prices. [7] used trading data from 2017 for the top 100 stocks on the Shenzhen Stock Exchange in 2017 which amounted to almost 10 million transactions. Using recall, precision and accuracy to evaluate the classification of price changes, they showed that TCNs outperformed both GARCH and LSTM baseline models.

2.4 Summary

To date, the most promising method used to predict gold prices has been the CNN-LSTM hybrid model [26, 37]. However, TCN is a framework adapted from CNNs for modelling sequential data [2] and has been found to have performed well in comparison to LSTMs [2, 24].

3 Methodologies

This paper implements a TCN on gold prices to determine how it compares to competing modelling approaches such as ARIMA, CNN, LSTM and a hybrid CNN-LSTM. In this section we initially give a description of the data used in the experimental analysis and following this briefly outline the methodology behind the chosen machine learning networks.

3.1 Dataset

The dataset used for training and testing the models in this study was obtained from Yahoo Finance via the yfinance library in the Python environment [11]. Specifically, the data was retrieved for Gold Futures using the symbol GC=F, between the 30th August 2000 and 31st December 2020. The data was then partitioned using a 80/20 split into training and test datasets. This meant the training dataset contained 4055 data points from 30th August, 2000 to 2nd December 2016 and the test set contained 1014 data points from 5th December 2016 to 31st December 2020. There was a further separation of the training set with an 80/20 split for train/validation, with the validation set used to tune the hyperparameters of all the models. To run the models effectively for a neural network, the training set was transformed using z-score normalisation and the time lag used was 64 days for each prediction.

3.2 Baseline Methods

ARIMA is a statistical model [4] that has been used extensively in time series studies. ARIMA models use a combination of autoregression, differencing and moving averages to forecast future values.

CNNs have been used extensively in image classification and have achieved improved results over the State of the Art (SOA) in comparison to previous research [23]. CNNs are able to achieve translational and scale invariance by learning from surrounding data points using convolutions and pooling techniques.

This concept has been applied to time series analysis as in [17, 40]. In addition, regularisation techniques such as Dropout [35] and Batch Normalization [18] have been developed to reduce over-fitting, resulting in better predictions when applied to out-of-sample data.

LSTMs were first proposed in 1995 by Hochreiter and Schmidhuber, but were not used extensively until machine learning began to gain traction. LSTMs are a special type of Recurrent Neural Network (RNN). While RNNs have memory that can be used to model sequence data they suffer from vanishing/exploding gradient problem [16]. LSTMs avoid this issue by creating gates: Input, Output and Forget. The Input gate decides which inputs should be allowed to propagate through the model by using a sigmoid function. The Forget gate determines which values should be dropped and finally the Output gate decides which values should be moved to the next time step. LSTMs have now gained wide popularity for sequential data modelling and have been used extensively for language modelling [39], translation [28], speech recognition [12] and time series analysis [10].

Using a CNN to extract features and feeding them to a LSTM model has proven to be a more powerful approach [26, 37]. The benefit of this hybrid approach, is to use the power of CNNs to extract features, which can then be fed sequentially or in parallel to a LSTM, thus focusing on the temporal features of the data elements.

3.3 Temporal Convolutional Networks

More recently, the concept of the Temporal Convolutional Network (TCN) was introduced by Lea et al. [24]. Bai et al. built upon this concept for sequence modelling [2], and is the approach used in this paper. They brought together three key concepts that define a TCN: Causal Convolutions, Dilated Convolutions and Residual Connections.

A causal convolution will only use current and prior data to produce its output. This means that the prediction at time t is only dependent on data elements up to time t as in Fig. 1. This shows the kernel can only see present and past values and not any future values.

Causal convolutions are not sufficient, by themselves, to successfully process long sequences. To achieve that, dilated convolutions are used to increase the size of the receptive field. When the dilation factor is one, this is equivalent to a regular convolution, however increasing the dilation factor to say two, will force the filter to only perform calculations on every second element. This enables the receptive field to be increased exponentially so the output is dependent upon a much larger receptive field, which reflects a larger set of previous time periods. TCNs use stacked convolutions with an increasing dilation factor, as in Fig. 1. This ensures the filter can hit each historical data element being processed, and also enables models to handle much longer historical data inputs. The experiments in this paper used a kernel size of 3 with 64 filters.

Residual connections were introduced by He et al. [14] and produced state of the art results on the ImageNet dataset [8]. Previously, deep networks architectures were confronted with vanishing/exploding gradients. However, residual

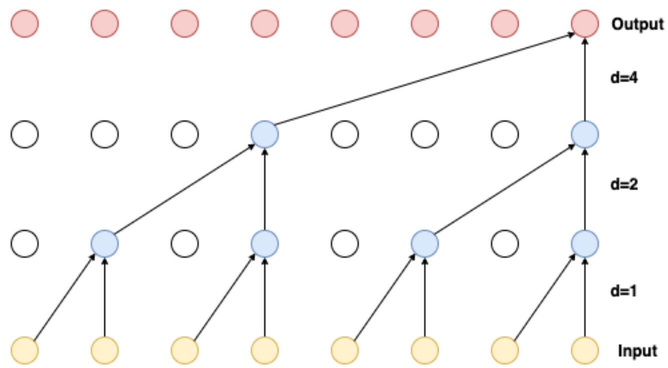

Fig. 1: Building blocks for a TCN:Dilated Causal Convolution

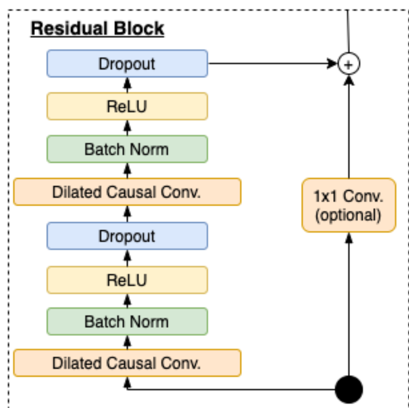

Fig. 2: Building blocks for a TCN:Residual Block

connections circumvent this issue by allowing the network to skip layers while passing the identity matrix in accordance with the formula in (1), where $\mathcal{H}(x)$ is the desired mapping, $\mathcal{F}(x)$ are the non-linear transformations and x is the identity mapping.

$$\mathcal{H}(x) = \mathcal{F}(x) + x \tag{1}$$

As the receptive field of a TCN is dependent on the depth of the network, it becomes essential to create a deeper network in order to achieve a larger receptive field. This paper implements residual connections, as in Fig. 2, using "Residual Blocks" with dilation factors of 1,2,4,8,16,32. The residual blocks use optional 1x1 convolutions which pass the input and are then added to the transformations using element-wise addition. These residual blocks can then be stacked with increasing dilation factors to achieve large receptive fields using a deep network.

To summarise, there are three main advantages of the TCN network [2]. Firstly, it allows data to be processed in parallel, as opposed to LSTMs which require input to be unfolded and processed step-by-step. Secondly, in comparison to LSTMs, TCNs provide a more stable gradient. Finally, the ability to control the dilation factor, kernel size and layer depth provides a lot of control over the receptive field, along with the ability for the output to be dependent on longer data elements.

One of the potential disadvantages of a TCN is they need to input the entire length of the history vector, which could be memory intensive, in contrast to LSTMs which only store a fixed length hidden state.

4 Results

In this section a univariate comparative analysis between a number of TCN architectures and four baseline approaches is given. Each approach attempted to predict the next days price on a given univariate series of gold prices.The baseline models selected were ARIMA, CNN, LSTM and the hybrid CNN-LSTM.

All models were implemented using the statsmodels library for ARIMA [33] and the machine learning approaches were implemented using the keras [5] and sci-kit learn libraries [32]. The metrics used to compare the success of the models were the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and R-Squared. For the formulas in (2) and (3) the number of predictions is represented by N, with p_i representing the predicted value and a_i being the actual value of the *i-th* instance. For the R-squared formula in (4), *RSS* represents the sum of square of the residuals and *TSS* is the total sum of squares.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|p_i - a_i| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i - a_i)^2} \tag{3}$$

$$R^2 = 1 - \frac{RSS}{TSS} \tag{4}$$

It's clear from the results in Table 1 that the TCN approach significantly outperformed all other baseline models with the TCN model producing a RMSE of 15.26, MAE of 10.05 and R^2 of 0.9954. The next best method was the hybrid CNN-LSTM model producing a RMSE of 23.53, MAE of 15.40 and R^2 of 0.9892, demonstrating that the inclusion of the CNN as a feature extractor does reduce the error over a stand-alone LSTM. However, it still falls short of the performance of the TCN in this scenario. The LSTM, CNN and ARIMA models produced RMSE scores of 27.39, 87.07 and 219.54 respectively.

The two best performing models, TCN and CNN-LSTM, were compared using walk-forward cross validation. Five train/test datasets were built with each

Table 1: Model performance on Gold Price prediction against the test dataset

	ARIMA	CNN	LSTM	CNN-LSTM	TCN
RMSE	219.54	87.07	27.39	23.53	15.26
MAE	147.69	76	19.5	15.4	10.05
R^2	0.0770	0.8534	0.9855	0.9892	0.9954

Table 2: Walk-Forward Cross Validation comparison between CNN-LSTM and TCN models

	TCN			CNN-LSTM		
Run	RMSE	MAE	R^2	RMSE	MAE	R^2
1	123.66	107.65	-0.3959	140.88	106.26	-0.8118
2	71.69	49.66	0.8401	439.36	366.79	-5.004
3	198.67	181.36	-0.3017	148.45	118.43	0.2732
4	12.03	17.4	0.9702	12.5	9.51	0.9678
5	26.26	17.4	0.9872	29.08	18.77	0.9842

training set increasing in size for subsequent folds using the TimeSeriesSplit library from sci-kit learn [32]. The results, shown in Table 2, provide further evidence of the TCN model outperforming the CNN-LSTM.

A chart of the comparison between the ground truth and the predictions from the TCN model, which was the best performing model, can be seen in Fig. 3. One thing to note is that all of the baseline models found it difficult to predict the up-swing in prices towards the end of the series in 2020, which is likely where most of the errors resulted. However, the TCN model was able to provide more accurate predictions for this sector of the test dataset.

One further benefit, that became apparent when training the models, was the CNN-LSTM model was much harder to optimize than the TCN. The amount of time required to perform hyper-parameter tuning was extensive in comparison to the TCN models. Also, the CNN-LSTM needed to run approximately 1000 epochs, but the TCN produced its best results in under 100 epochs. Both of these points are important when considering both computational power and time required to produce a model.

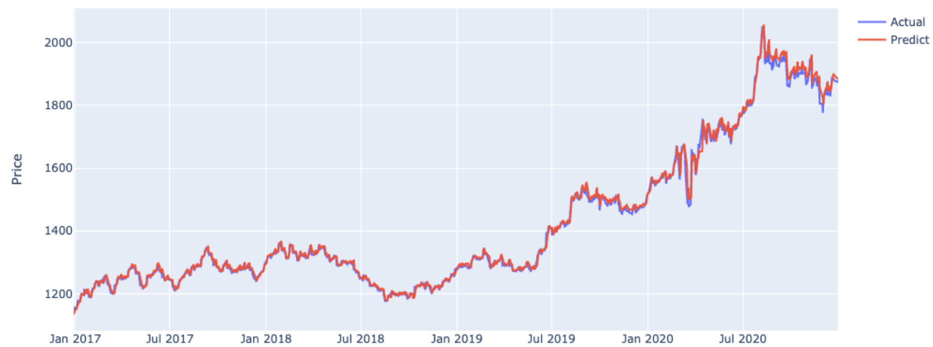

Fig. 3: Predictions versus Actual for the TCN model against the test dataset

5 Conclusion

This research undertook a comparative analysis between Temporal Convolutional Networks (TCNs), the current state of the art machine learning approaches and a traditional time series model in gold price prediction. The results demonstrated that the TCNs consistently outperformed the other approaches chosen in this study and reduced the error by more 27% in comparison with the best preforming non-TCN approach.

While Gold has traditionally been the chosen safe-haven commodity, recent market behaviour throughout the pandemic has indicated that crypto-currencies have become potential alternates. While crypto-currency prediction has been extensively studied using deep learning approaches, understanding the relationship with gold could potentially have an import impact on future predictions. While this study focused on the performance of TCNs on univariate prediction and future work would possibly benefit from the inclusion of exogenous variables such as crypto-currencies,stock market indices and economic indicators [29].

References

1. Alameer, Z., Abd Elaziz, M., Ewees, A.A., Ye, H., Jianhua, Z.: Forecasting gold price fluctuations using improved multilayer perceptron neural network and whale optimization algorithm. Resources Policy **61**, 250–260 (2019)
2. Bai, S., Kolter, J.Z., Koltun, V.: An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint arXiv:1803.01271 (2018)
3. Balcilar, M., Gupta, R., Pierdzioch, C.: Does uncertainty move the gold price? new evidence from a nonparametric causality-in-quantiles test. Resources Policy **49**, 74–80 (2016)
4. Box, G.E., Jenkins, G.M., Reinsel, G.C., Ljung, G.M.: Time series analysis: forecasting and control. John Wiley & Sons (2015)
5. Chollet, F., et al.: Keras (2015), https://github.com/fchollet/keras

6. Corti, C.W., Holliday, R.J.: Commercial aspects of gold applications: from materials science to chemical science. Gold Bulletin **37**(1), 20–26 (2004)
7. Dai, W., An, Y., Long, W.: Price change prediction of ultra high frequency financial data based on temporal convolutional network. arXiv preprint arXiv:2107.00261 (2021)
8. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 248–255. Ieee (2009)
9. Fang, L., Chen, B., Yu, H., Qian, Y.: The importance of global economic policy uncertainty in predicting gold futures market volatility: A garch-midas approach. Journal of Futures Markets **38**(3), 413–422 (2018)
10. Filonov, P., Lavrentyev, A., Vorontsov, A.: Multivariate industrial time series with cyber-attack simulation: Fault detection using an lstm-based predictive data model. arXiv preprint arXiv:1612.06676 (2016)
11. Finance, Y.: Gold price dataset from yahoo finance. https://finance.yahoo.com/quote/GC=F?p=GC=F&.tsrc=fin-srch, accessed: 8th Feb, 2021
12. Graves, A., Jaitly, N., Mohamed, A.r.: Hybrid speech recognition with deep bidirectional lstm. In: 2013 IEEE workshop on automatic speech recognition and understanding. pp. 273–278. IEEE (2013)
13. Guha, B., Bandyopadhyay, G.: Gold price forecasting using arima model. Journal of Advanced Management Science **4**(2) (2016)
14. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
15. He, Z., Zhou, J., Dai, H.N., Wang, H.: 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 (2019)
16. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation **9**(8), 1735–1780 (1997). https://doi.org/10.1162/neco.1997.9.8.1735
17. Hoseinzade, E., Haratizadeh, S.: Cnnpred: Cnn-based stock market prediction using a diverse set of variables. Expert Systems with Applications **129**, 273–285 (2019)
18. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: International conference on machine learning. pp. 448–456. PMLR (2015)
19. Iqbal, J.: Does gold hedge stock market, inflation and exchange rate risks? an econometric investigation. International Review of Economics & Finance **48**, 1–17 (2017)
20. Jian-Hui, Y., Wei, D.: Prediction of gold price based on wt-svr and emd-svr model. In: 2012 Eighth International Conference on Computational Intelligence and Security. pp. 415–419. IEEE (2012)
21. Jianwei, E., Ye, J., Jin, H.: A novel hybrid model on the prediction of time series and its application for the gold price analysis and forecasting. Physica A: Statistical Mechanics and its Applications **527**, 121454 (2019)
22. Kim, T., Kim, H.Y.: Forecasting stock prices with a feature fusion lstm-cnn model using different representations of the same data. PloS one **14**(2), e0212320 (2019)

23. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems **25**, 1097–1105 (2012)
24. Lea, C., Flynn, M.D., Vidal, R., Reiter, A., Hager, G.D.: Temporal convolutional networks for action segmentation and detection. In: proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 156–165 (2017)
25. Liu, D., Li, Z.: Gold price forecasting and related influence factors analysis based on random forest. In: Proceedings of the Tenth International Conference on Management Science and Engineering Management. pp. 711–723. Springer (2017)
26. Livieris, I.E., Pintelas, E., Pintelas, P.: A cnn–lstm model for gold price time-series forecasting. Neural Computing and Applications pp. 1–10 (2020)
27. Lucey, B.M., Sharma, S.S., Vigne, S.A.: Gold and inflation (s)–a time-varying relationship. Economic Modelling **67**, 88–101 (2017)
28. Luong, M.T., Sutskever, I., Le, Q.V., Vinyals, O., Zaremba, W.: Addressing the rare word problem in neural machine translation. arXiv preprint arXiv:1410.8206 (2014)
29. Mariana, C.D., Ekaputra, I.A., Husodo, Z.A.: Are bitcoin and ethereum safe-havens for stocks during the covid-19 pandemic? Finance research letters **38**, 101798 (2021)
30. Mombeini, H., Yazdani-Chamzini, A.: Modeling gold price via artificial neural network. Journal of Economics, business and Management **3**(7), 699–703 (2015)
31. Oord, A.v.d., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A., Kavukcuoglu, K.: Wavenet: A generative model for raw audio. arXiv preprint arXiv:1609.03499 (2016)
32. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research **12**, 2825–2830 (2011)
33. Seabold, S., Perktold, J.: statsmodels: Econometric and statistical modeling with python. In: 9th Python in Science Conference (2010)
34. Siami-Namini, S., Tavakoli, N., Namin, A.S.: A comparison of arima and lstm in forecasting time series. In: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). pp. 1394–1401. IEEE (2018)
35. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research **15**(1), 1929–1958 (2014)
36. Verma, S., Thampi, G., Rao, M.: Ann based method for improving gold price forecasting accuracy through modified gradient descent methods. IAES International Journal of Artificial Intelligence **9**(1), 46 (2020)
37. Vidal, A., Kristjanpoller, W.: Gold volatility prediction using a cnn-lstm approach. Expert Systems with Applications **157**, 113481 (2020)
38. Yousef, I., Shehadeh, E., et al.: The impact of the covid-19 on gold price volatility. Int J Econ Bus Adm **8**(4), 353–64 (2020)
39. Zaremba, W., Sutskever, I., Vinyals, O.: Recurrent neural network regularization. arXiv preprint arXiv:1409.2329 (2014)
40. Zhao, B., Lu, H., Chen, S., Liu, J., Wu, D.: Convolutional neural networks for time series classification. Journal of Systems Engineering and Electronics **28**(1), 162–169 (2017)