TensorFlow/Keras versus H2O, Round 2: Predicting Currency Prices

Kenneth E. Davis M.S.

Abstract (#318427 JSM 2021)

In Round 1 we examined the prediction of the SP500 index with machine learning methods using TensorFlow in Python and H2O in R. Although both performed nearly identically in predicting prices over time, H2O in R was found to confer better loss protection in volatile conditions by a slim margin. Will this relationship hold true in the currency market.

The present study will continue the series by examining TensorFlow in Python and H2O in R for predicting currency prices in Round 2. The results of the EURUSD currency pair will be analyzed in the pre, intra, and post pandemic period, and analysis characteristics between Rounds 1 and 2 compared.

Differences between problem types, or financial markets, while seeking the identical outcome of profitability and conservation of principle will be described. We will evaluate the effectiveness of open source machine learning tools using workstation CPU, and examine the story expressed by the data, both the relationship between predictors and outcome, and the relationships between the explanatory variables themselves when constructing new machine learning models in a continuing variety of financial markets.

With currencies, TensorFlow/Keras in Python was found to outperform H2O in R for total return, reporting +77.7% versus +70.3% respectively. Further, TensorFlow/Keras was found to confer better loss protection at -4.6% versus -14.1% respectively. This corresponds to a maximum loss of -23.2% versus -70.6% respectively at 25 times currency leverage in the 2 year duration of study. Leveraged gains corresponded to +388.5% and +351.8% respectively. The correct call rate was 58% for both platforms. Despite mild differences, both platforms were shown to be quite adequate in developing successful strategies for prediction in the currency market.

KeyWords: Neural Network, Machine Learning, TensorFlow Keras, H2O, Currencies, Financial Markets

TensorFlow/Keras versus H2O, Round 2: Predicting Currency Prices

Kenneth E. Davis M.S.

1. Introduction

Considerable achievements in data science have been reported in recent years[1-10] along with a rapid evolution of machine learning tools. Still, there has remained the open question of if similar advancements can be made in the financial markets.

Prior to the last five years, the financial markets have remained an area with comparatively few openly published admissions of actionable success.[11-12] Under reporting of successful machine learning strategies has likely been due to the idea that the more widespread a successful investing methodology becomes, the less likely it is to be profitable above that of the overall market. Despite these conditions of implied secrecy, ready accessibility to machine learning tools as well as large quantities of publically available data make prediction in investable markets a tangible possibility, even when solving difficult or secretive analysis problems. Further, newly emerging methodologies and model complexities[13-14] may make such dire protection of successful methods less necessary.

The present analysis will examine prediction in the currency market, in this case the EUR/USD currency pair, using widely available, open source machine learning tools along with easily available financial data provided cost-free on the web. We will compare TensorFlow/Keras in Python to H2O in R for predicting currency returns, and examine the differences between foreign exchange (forex) prediction and SP-500 stock market prediction in this continuing series published at JSM [15-16]. Successful (and unsuccessful) performance, including the volatile Covid-19 pandemic period, will be examined in terms of profitability, design and methodology, and protection of principle.

2. Materials and Methods

2.1 Software and OS

Analysis Software
Python 3.6.8 with Spyder 3.3.2
TensorFlow 1.6.0 (for cross compatibility with Amazon Cloud installations)
R Studio 1.456
R 3.5.0
Anaconda Navigator 1.9.6

OS and CPU
Operating System: Windows 10 Pro
CPU: Intel Core i7 8-Core CPU at 3.4 GHz

2.2 Data Sources

a.) EUR/USD historical data Yahoo Finance Dec 12, 2005 thru June 7, 2021
b.) FXE historical data Yahoo Finance Dec 12, 2005 thru June 7, 2021

2.3 Study Duration

Training Period: Dec 2005- May 2019
Validation Period: None
Test Period: June 2019 -May 2021

2.4 Base Model[17]

Outcome Variable[17]:

1. 10-day future return percent

Predictors[17]

1. 14-day Moving Average (MA)
2. 14-day Relative Strength Index (RSI)
3. 200-day Moving Average
4. 200-day Relative Strength Index
5. 5-day prior return percent

2.5 Additional Variables Considered

Additional Outcome Variables

1. 1-day future return percent
2. 3-day future return percent
3. 5-day future return percent

Additional Predictors

6. 30-day Moving Average
7. 30-day Relative Strength Index
8. 50-day Moving Average
9. 50-day Relative Strength Index

2.6 Model Architecture

The initial base architecture[17] and settings were:
2- hidden layers
50 node in layer 1
10 node in layer 2
1 output layer
50 epochs

Final architectures were:
TensorFlow/Keras in Python.
2- hidden layers with
100 node in layer 1
20 node in layer 2
1 output layer
750 epochs

H2O in R.
2-hidden layers with
100 nodes in layer 1
20 nodes in layer 2
1 output layer
10,000 epochs

Common Settings were:
Sequential Model
ReLU Activation
1-node output layer with Linear Activation
Adam Optimizer
Loss Function of MSE (TensorFlow/Keras), and RMSE (H2O).

2.7 Final output signal

a.) BUY if the neural network prediction value was positive
b.) SELL if the neural network prediction value was negative
c.) prediction of 0 did not occur in the study duration

2.8 Final Model and Analysis Detail

All models, parameters and settings were constructed at the study's initiation in 2019, however model selection from the n=35 model loop in each platform did change for the purpose reporting this analysis. The final model predictors and outcome were unchanged across the study duration, and were:

<u>Final Outcome Variable:</u>
1. 5-day future return percent

<u>Final Predictors:</u>

2. 14-day Moving Average
3. 14-day Relative Strength Index
4. 30-day Moving Average
5. 30-day Relative Strength Index
6. 50-day Moving Average
7. 50-day Relative Strength Index
8. 200-day Moving Average
9. 200-day Relative Strength Index
10. 5-day prior return percent

Cumulative Returns were the cumulative daily sum of one position's gain or loss resulting from each BUY or SELL signal from the model at the end of trading hours (market close) in New York City (4:00 p.m. eastern standard time) . The position was considered closed at the market close of the 5^{th} day. BUY signals were designed to be executed in a buy account, and SELL signals were considered executed in a separate sell account.

Leveraged Cumulative Returns were the cumulative sum of each positions gain or loss based on 1/5 of total principle per day, multiplied by the leverage ratio. 25 times leverage values were therefore calculated as the cumulative 5-day return times five (=25/5).

2.9 Definition of Pre, Intra, and Post Pandemic Period

Training data date range was: 12/12/2005 (yahoo finance start date for FXE) thru 05/31/2019.

Pre pandemic was defined as: 6/1/2109 thru 12/31/2019
Intra Pandemic was defined arbitrarily as: 1/1/2020 thru 5/31/2020
Post Pandemic was defined as: 6/1/2020 thru 5/31/2021,with the final day of 6/7/2021 set to obtain the final 5-day future return

The U.S. Senate passed the Covid-19 economic stimulus package on March 25, 2020.

3. Results

Validation Period: None
Test Period (actual): 06/03/2019 - 05/28/2021

3.1 Proportion Correct Calls (buy or sell)

Proportion Correct (n=503 days)			
	% correct	proportion	p-value^
TensorFlow/Keras Python	58.3%	293/503	0.0001
H2O gnu R	58.4%	294/503	<0.0001

^binomial test versus .50 null

3.2 Overall Percent Return and Maximum Loss

Cumulative Model Returns		
	% cum gain	max loss
TensorFlow/Keras Python	+77.7%	-4.6%
H2O gnu R	+70.3%	-14.1%

3.3 Overall Market Return and Maximum Loss

Overall Market June 2019 thru May 2021		
	return	max loss
EUR/USD market	+6.5%	-5.1%
SP-500	+53.2%	-33.7%

based on closing prices

3.4 Leveraged Percent Return and Maximum Loss

Leveraged 25x Returns		
	% cum gain	max loss
TensorFlow/Keras Python	+388.5%	-23.2%
H2O gnu R	+351.8%	-70.6%

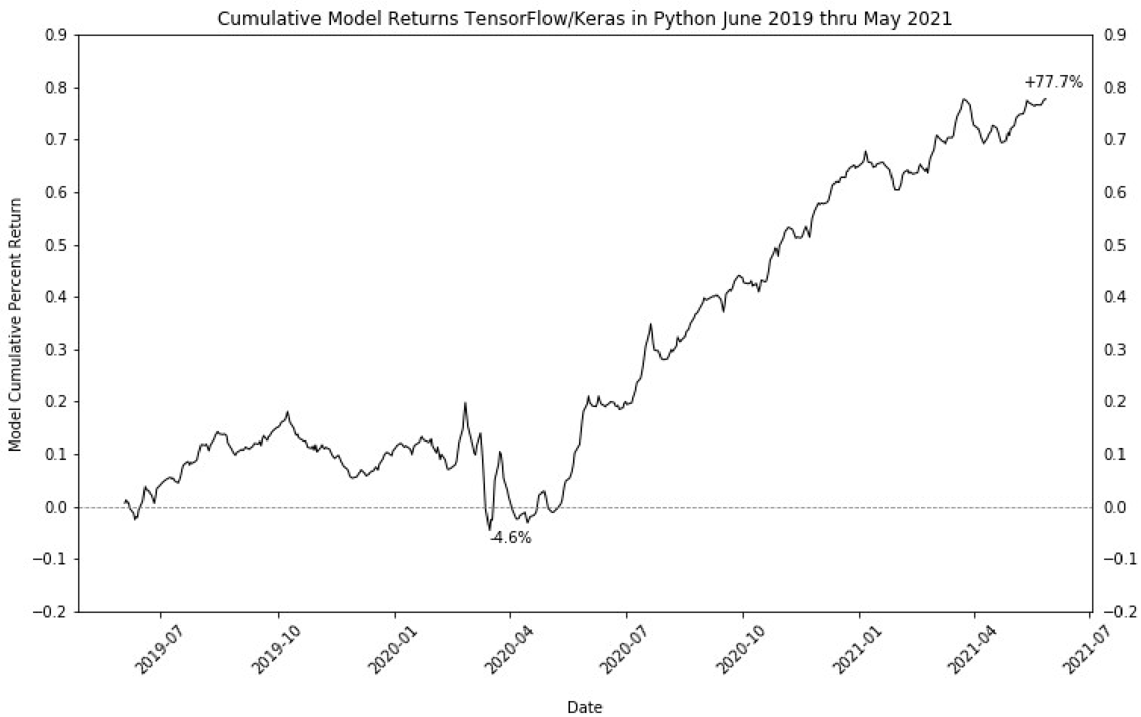
3.5 Percent Correct Calls by Time Period

	TF	H2O
Pre Pandemic	56%	58%
Intra Pandemic	56%	45%
Post Pandemic	60%	64%
Total	58%	58%

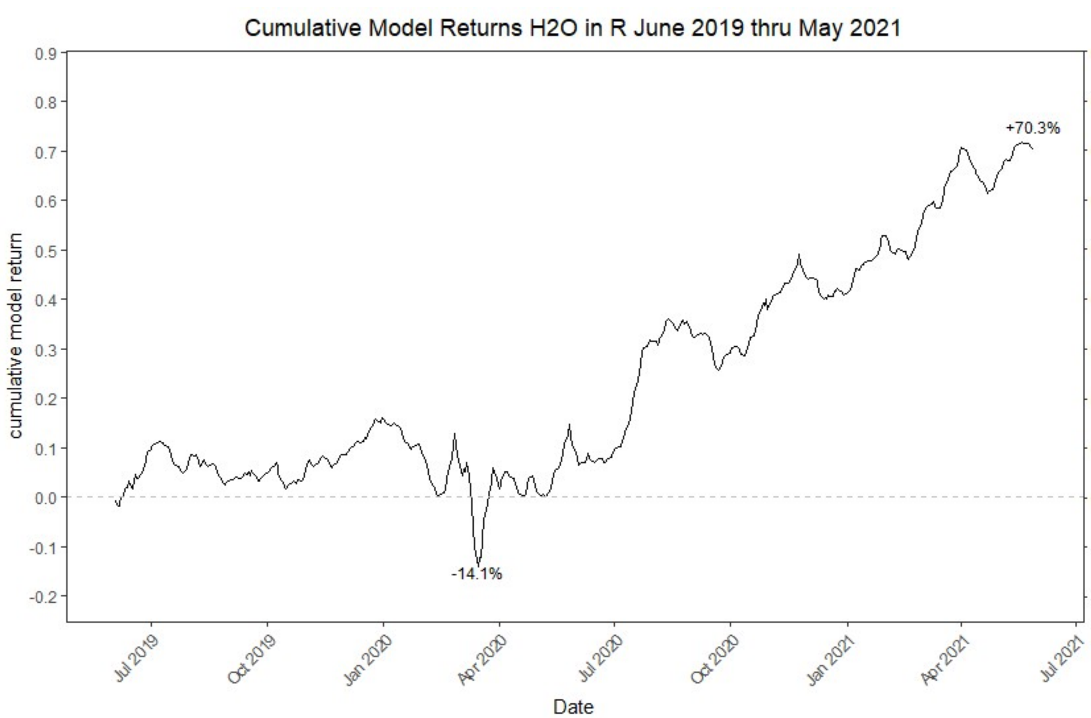
3.6 Percent Return by Time Period

	TF	H2O
Pre Pandemic	+10.6%	+16.0%
Intra Pandemic	+7.5%	-5.3%
Post Pandemic	+59.6%	+59.6%
Total	+77.7%	+70.3%

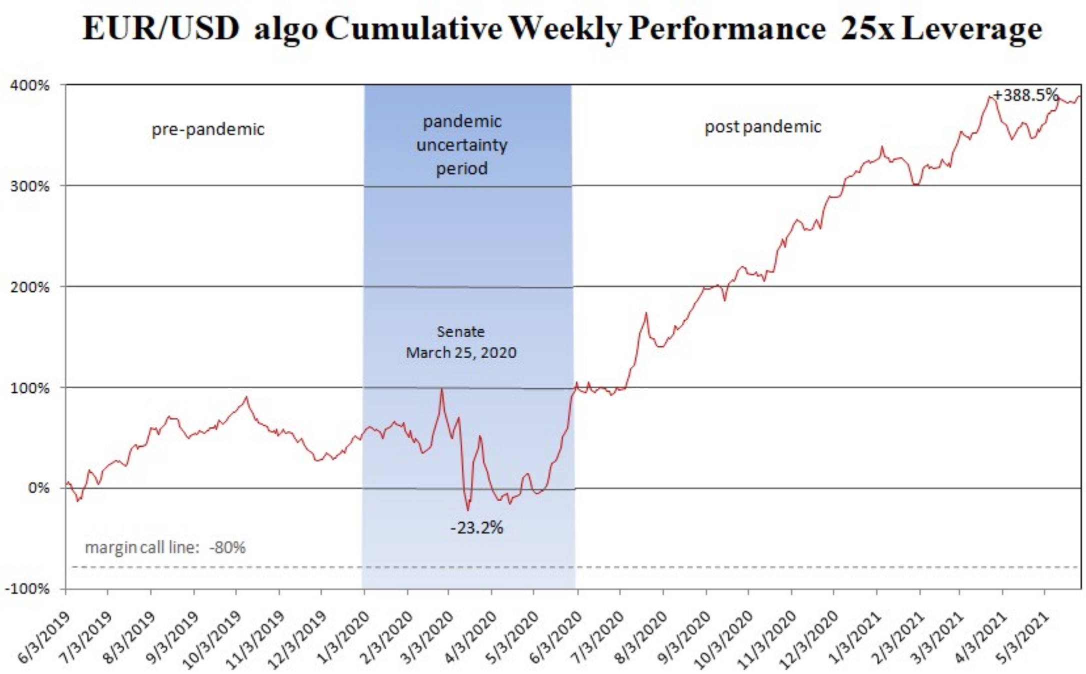
3.7 Cumulative Returns, TensorFlow/Keras Python 06/01/2019 - 05/31/2021



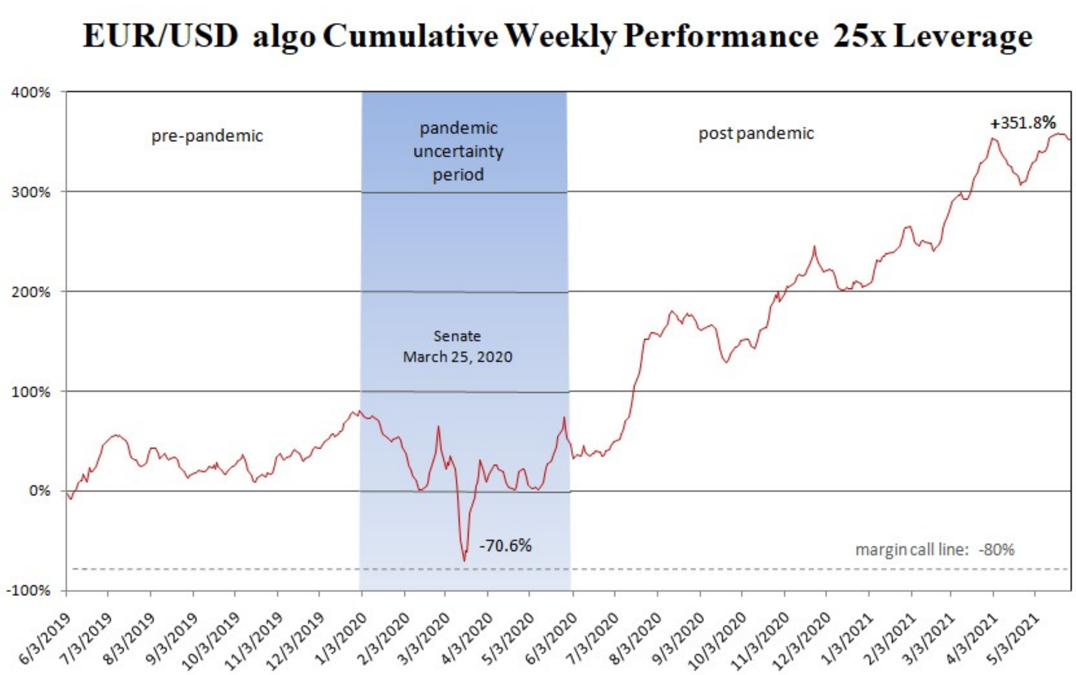
3.8 Cumulative Returns H2O R 06/01/2019 - 05/31/2021



3.9 Leveraged Cumulative Returns TensorFlow/Keras Python 06/01/2019 - 05/31/2021



3.10 Leveraged Cumulative Returns H2O R 06/01/2019 - 05/31/2021



3.11 Multicollinearity / Variance Inflation Factor (VIF)

Predictor	VIF^
rsi30	1724
rsi50	1328
rsi14	170
ma50	46
ma200	44
ma30	26
rsi200	23
ma14	11
5d_close_pct	4

^VIF>10 is high

4. Discussion

4.1 Current Findings

Prediction in the financial markets has been increasingly reported using feed-forward neural network[18,19], recurrent neural network/LSTM[13,14, 19-22], hidden Markov model[13,24], and ensemble methods,[19] among others.

The present analysis reports a 58% accuracy rate for both TensorFlow/Keras in Python and H2O in R, using feed forward neural network and technical indicators performed on an 8-core workstation CPU. Total gain over the 2-year outcome duration including the volatile pandemic period, was +77.7% for TensorFlow/Keras in Python, with a buy or sell signal generated on 100% of trading days. This corresponds to a 2-year return of +388.5% when using 25 times currency account leverage. For TensorFlow/Keras in Python, the maximum account loss was -4.6%, or -23.2% at 25 times leverage, avoiding the leveraged margin call limit with a relatively large degree of latitude. Comparably, H2O in R reported a gain of +70.3% in the study duration, corresponding to +351.8% in the leveraged account, with a maximum loss of -14.1% unleveraged, and a relatively high -70.3% if using 25 times leverage.

For proportion of correct calls specifically, TensorFlow/Keras in Python reported a 58.3% correct call rate, and H2O in R a nearly identical 58.4% correct call rate in the outcome duration of the study.

4.2 Comparisons to the Current Literature

Signal accuracies (e.g. a successful ‘buy’ or ‘sell’ signal) between 52%-59% for machine learning methods in various time frames and financial markets have been reported in the current literature[13,14,21,22,26]. Furthermore, signal accuracy rates of 70%[13] and 80%[14] or higher [25,26] have also been reported, usually with increasing model complexity and/or greater post signal processing. In the current publication year, Yıldırım et al.[13] reported a 73.6% correct calls, triggered on 40.4% of the available trading days. In the Yıldırım analysis, forex markets were analyzed using both technical and macroeconomic predictors in two separate but later combined signals, which ultimately indicated buy, sell, or no action, in holding periods of 1 to 5days. Hu[25] reported an 89.0% correct call rate for the Dow Jones Industrial Average with predictors that included the use of Google Trends data.

Comparatively fewer studies quantify percentage gains over longer time periods specifically, however some have reported returns in time frames of months[21] and annually[13, 23]. Rigorous comparisons of machine learning methods, model types, and innovative processing structures are the focus of many published works, with more works citing increased success with LSTM[13,14,19-22, 26].

4.3 Predictor Selection, Multicollinearity

It is well known that multicollinearity is not a problem for machine learning methods, and non-linearity is easily handled as well. Consequently, direct quantification of correlation and multicollinearity among predictors is rarely necessary to report[27]. The current analysis however reported an alarming rate of variance inflation of up to 1724, where values over 10 are certainly of concern in linear regression models. In the present analysis, additional predictors did not appear to jeopardize the neural network model, even though, for example, relative strength at 30 days is highly correlated to relative strength at 50 days (not directly reported, but visualized in subsection 3.11 above). For model building, it may be better to have more predictors than less even if they are similar, and better performing models in the current literature tended to have rather higher numbers of predictors than fewer[13].

4.4 Comparison to the SP500 Index Model

There were notable analysis differences from the SP500 Index model in previous work[15.] For the SP500 model, out of range characteristics of test data variables and the scaling of values not found in training data appeared to contribute to increased variance and greater model inefficiencies. For the SP500 model, the maximum losses were -4.3% and -3.9% for TensorFlow/Keras in Python and H2O in R, respectively, however the overall market loss in that duration was only- 3.6%. In this way, there was little or no protection of principle in the SP-500 models.

In the present work, technical indicator values occurring in the training data covered the entire range of values in the test set in almost every case (not reported), and the values of currencies (as well as 5-day return percentages) appeared to oscillate within a similar, repeating set of ranges. Preservation of capital appeared to be much more easily obtained, with the greatest loss at -4.6% in the pandemic period (for TensorFlow/Keras in Python), compared to a loss of -33.7% in the SP500 Index at the pandemic low.

4.5 Importance of Preservation of Capital for Currencies

Preservation of capital is important in currencies due to the real possibility of a substantial, nearly complete margin call on the account when leveraged. Although trading costs for currencies are arguably negligible, with buy/sell spreads of less 0.02% and interest rates of less than 2% annually at present, substantial losses in leveraged conditions certainly involve higher risk. In the leveraged R model, -70.6% would take us very close to the margin call limit (-80%, for example at OANDA), threatening to mandatorily/automatically close all account positions in the absence of greater added capital. Consequently, the present analysis would appear to indicate the use of a slightly limited leverage level, e.g. 25 times, even though greater leverage is available (i.e. 50 times leverage at OANDA and Interactive Brokers).

4.6 Model Stability, Degeneration, and Weaknesses of the Study

The primary weakness of this analysis is the arbitrary selection for models in the test period. While the H2O model in R was selected by the highest R^2 value in the training data, the best model in TensorFlow/Keras in Python was selected for feasibility solely due to performance in the test period. It would be quite easy to run several models and merely pick the best fit to the test period, so it is clear that better model selection rules need to be created, such as a well defined validation period and /or cross validation. Further, there is no assurance that any selected model would continue to perform indefinitely over time, regardless of whether a well defined validation period or holdout training was used. Interestingly, both models presented in the current analysis regained accuracy after the pandemic duration, and nearly identical accuracies between them were observed.

Other weakness include a wide range of unanalyzed technical indicators remaining unevaluated, and only one of many possible machine learning architectures examined (e.g. LSTM, Bayesian Neural Network, CNN, etc). Further, there were numerous arrays of potential hyper parameters left unexplored.

4.7 Future Applications, Gold, Silver and Oil

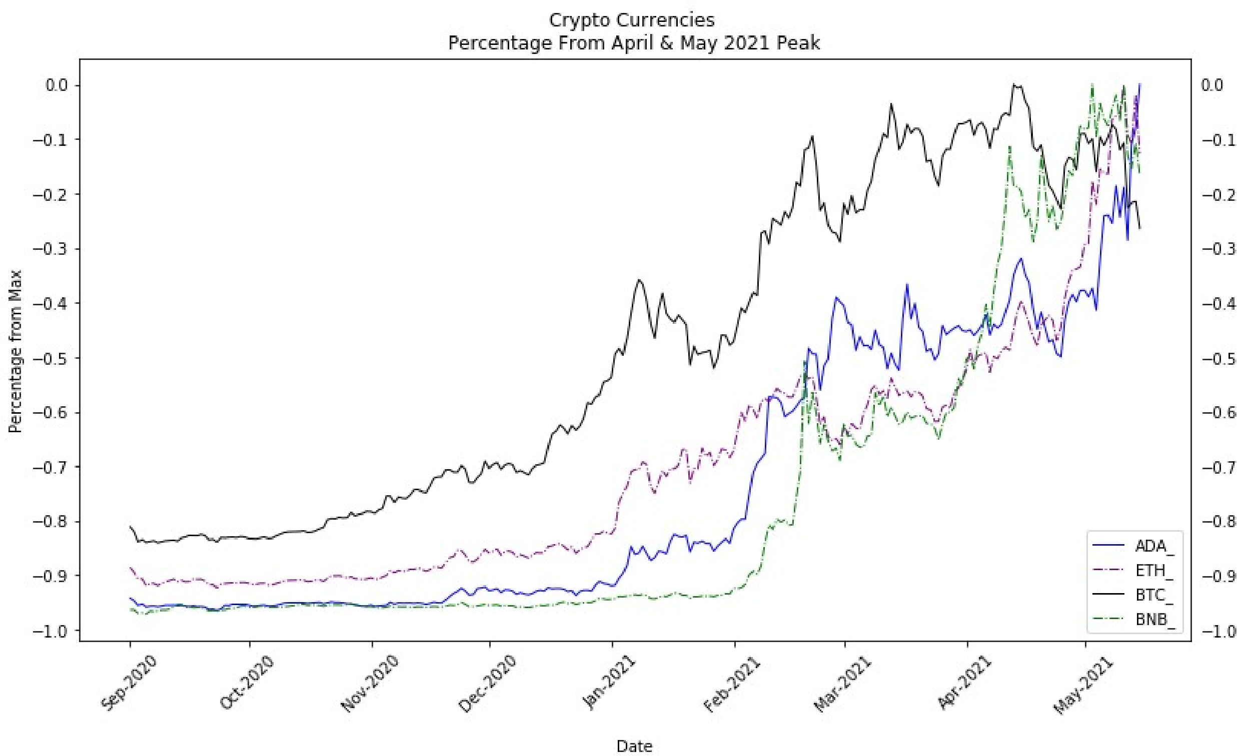
Forces affecting the prices in markets such as gold[28], silver, oil and other commodities are likely somewhat different from fiat currencies and the SP-500 Index. These differences include the changing policies of various governments and production decisions of such entities as OPEC. These would also include the overall effects of supply and demand involved with commodities, as well as interest rate differences between markets globally. In this way, commodities may represent markets with different predictor sets than the present analysis, and may represent interesting challenges for machine learning models and data sciences going forward.

Figure in 4.7: Example of gold and silver prices over time

4.8 Crypto Currencies[29-31]

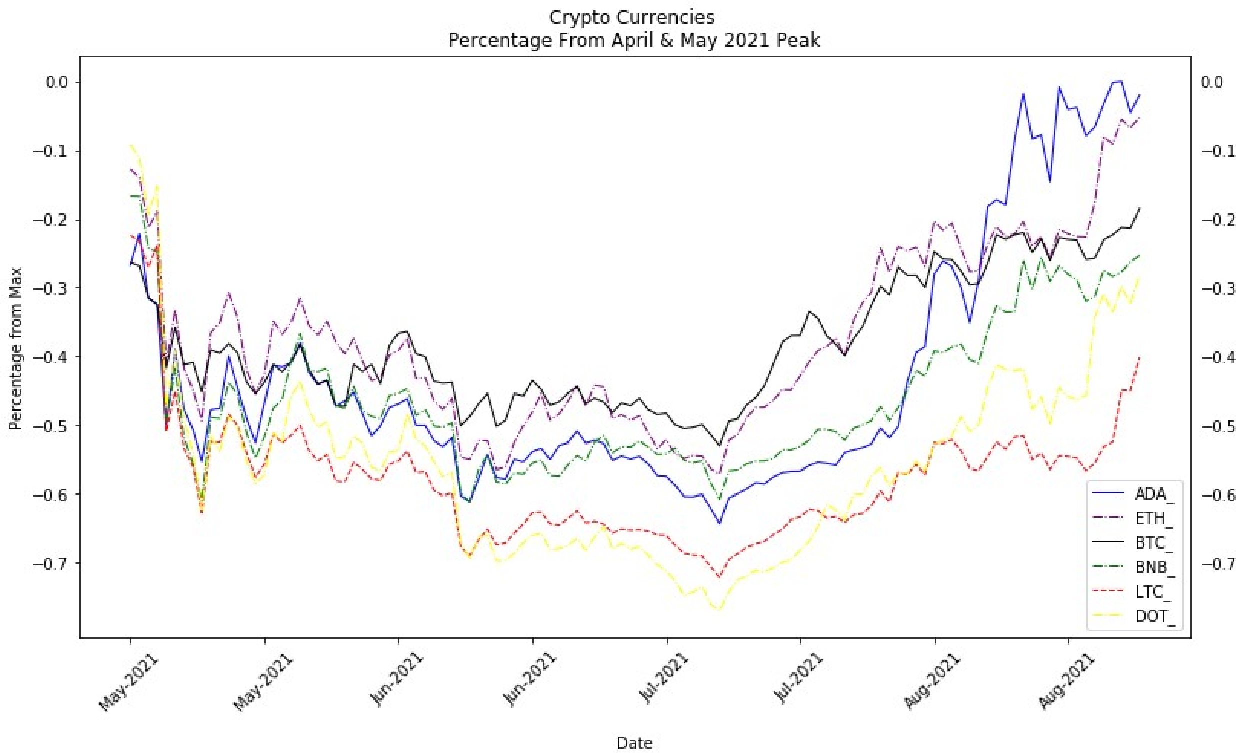
Differences in technical predictor sets are worth some consideration in newer markets. For example, it is unlikely that the 200-day moving average used in the present analysis would be as predictive for crypto currencies at the present time. The price increase of all the major crypto currencies beginning in late 2020 was relatively large as new eras of crypto currency investment began to appear.

Figure in 4.8a: Rapid increases in values may invalidate long term predictors



As these new markets continue to develop, dynamically changing predictor sets may likely be necessary. Such drastic shifts in prices give rise to interesting machine learning challenges (such as the use of reinforcement learning methods) in rapidly evolving markets such as the crypto currencies[29-31]. Further, wide price differences and change in price differences between markets may need to be viewed in different ways, such as percentage from peak rather than dollar value.

Figure in 4.8b: Crypto currency as percentage of peak, rather than dollar value



4.9 Conclusions

Both TensorFlow/Keras in Python and H2O in R produced tangible trading success for currencies, with models that performed clearly profitably and relatively similarly. Models based on technical indicators appear to be a viable, obtainable analysis tool in the trading of currencies, although more sophisticated models are now beginning to emerge with what may be greater success. Further, the current literature is beginning to identify more options and innovative strategies for many financial markets, and more studies are beginning to quantify longer term return percentages and return on capital[26]. Controlling for losses are shown to be possible in the present analysis, but remain under reported in the current literature.

Correspondence: kedavis@alumni.iu.edu

5. References

1. Silver D, Huang A, Maddison C J, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D. Mastering the game of Go with deep neural networks and tree search. Nature 529(7587), 484–489. 2016.

2. Brown N, Sandholm T. Baby Tartanian8: Winning agent from the 2016 annual computer poker competition. In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16), pages 4238–4239. 2016.

3. Dannenberg, R. B. Artificial intelligence, machine learning, and music understanding. Proceedings of the 2000 Brazilian Symposium on Computer Music: Arquivos do Simposio Brasileiro de Computao Musical. Brazil. 2000.

4. Kasten Y, Doktofsky D, Kovler I. End-to-end convolutional neural network for 3D reconstruction of knee bones from bi-planar X-ray images. In International Workshop on Machine Learning for Medical Image Reconstruction. pp. 123–133. Springer, 2020.

5. Elgammal A. AI Is Blurring the Definition of Artist. American Scientist 107, 1: 18. 2019

6. Devlin J, Chang M, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv:1810.04805. 2018.

7. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In Proceedings of NIPS, pages 1106–1114, 2012.

8. Silver D ,Hubert T Schrittwieser J, Antonoglou I, Lai M, Guez A, Lanctot M, Sifre L, Kumaran D, Graepel T, Lillicrap T, Simonyan K, Hassabis D. A general reinforcement learning algorithm that masters chess shogi and Go through self-play, Science, vol. 362, no. 6419, pp. 1140-1144. 2018.

9. Johnson M, Schuster M, Le QV, Krikun M, Wu Y, Chen Z, Thorat N, Viégas FB, Wattenberg M, Corrado G, Hughes M, Dean J. Google's multilingual neural machine translation system: Enabling zero-shot translation. CoRR abs/1611.04558. http://arxiv.org/abs/1611.04558. 2016.

10. Artificial intelligence is changing every aspect of war. Battle algorithm.
A new type of arms race could be on the cards. The Economist. September 7, 2019.

11. Cavalcante RC, Brasileiro RC, Souza VLF, Nobrega J, Oliveira ALI. Computational intelligence and financial markets: a survey and future directions. Expert Syst Appl 55:194–211. 2016.

12. Bahrammirzaee A. A comparative survey of artifcial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. Neural Comput Appl 19:1165–1195. 2010.

13. Yıldırım DC, Toroslu IH, Fiore U. Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators. Financ Innov 7:1. 2021

14. Fischer, T, Christopher K. Deep learning with long short-term memory networks for financial market predictions. (working paper) FAU Discussion Papers in Economics, No. 11/2017, Friedrich-Alexander-Universität Erlangen-Nürnberg Institute for Economics, Nürnberg. 2017.

15. Davis KE. TensorFlow versus H20: Predicting the SP500. 2019. In JSM Proceedings, Statistical Computing Section. Alexandria, VA: American Statistical Association.

16. Davis KE. Advantageous Statistical Tools for Stock Market Investing. 2018. In JSM Proceedings, Statistical Computing Section. Alexandria, VA: American Statistical Association.

17. George, Nathan. Regis University. https://www.datacamp.com/courses/machine-learning-for-finance-in-python

18. Qiu M, Song Y. Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model. PLoS ONE 11(5): e0155133. https://doi.org/10.1371/journal.pone.0155133. 2016.

19. Nabipour M, Nayyeri P, Jabani H, Mosavi A, Salwana E, S. Deep Learning for Stock Market Prediction. Entropy. 22(8):840. https://doi.org/10.3390/e22080840. 2020.

20. Galeshchuk S, Mukherjee S. Deep networks for predicting direction of change in foreign exchange rates. Intell Sys Acc Fin Mgmt. DOI 10.1002/isaf.1404. 2017.

21. Gao Q, Ze H (supervisor). Stock market forecasting using recurrent neural network: M.Sc. thesis presented to Faculty of Graduate School, University of Missouri-Columbia. USA. 2016.

22. Pang X, Zhou Y, Wang P, Chang V. An innovative neural network approach for stock market prediction. J Supercomput 76, 2098–2118. https://doi.org/10.1007/s11227-017-2228-y. 2020.

23. Ji Q, Neerchal NK. Creating Stock Portfolios Using Hidden Markov Models. 2019. In JSM Proceedings, Statistical Computing Section. Alexandria, VA: American Statistical Association.

24. Gupta A, Dhingra B. Stock market prediction using Hidden Markov Models, 2012 Students Conference on Engineering and Systems 2012 pp. 1-4, doi: 10.1109/SCES.2012.6199099.

25. Hu H, Tang L, Zhang S, Wang H. Predicting the direction of stock markets using optimized neural networks with Google Trends. Neurocomputing 285:188–195. 2018.

26. Hu Z, Zhao Y, Khishi M. A Survey of Forex and Stock Price Prediction Using Deep Learning (review). Appl. Syst. Innov. 2021,4,9. https://doi.org/10.3390/asi4010009. 2021.

27. Veaux RD, Ungar LH. Multicollinearity: A tale of two nonparametric regressions Selecting Models from Data. Artificial Intelligence and Statistics IV Volume 89. ISBN : 978-0-387-94281-0. 1994.

28. Shaikh S, Sharma S. Predicting Gold prices: Using Crude Oil, Currency movement & Stock Market. IOSR Journal of Economics and Finance (IOSR-JEF) e-ISSN: 2321-5933, p-ISSN: 2321-5925. Volume 12, Issue 3 Ser. II (May –June), pp 20-23. 2021.

29. Biswas S, Pawar M, Badole S, Galande N, Rathod S. Cryptocurrency Price Prediction Using Neural Networks and Deep Learning. 7th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 408-413, doi:10.1109/ICACCS51430.2021.9441872. 2021.

30. Yilmaz, KN, Hazar, BH. Predicting future cryptocurrency investment trends by conjoint analysis. Journal of Economics, Finance and Accounting (JEFA), V.5(4), p.321-330. 2018.

31. Prakask Kolla, B. Predicting Crypto Currency Prices Using Machine Learning and Deep Learning Techniques. International Journal of Advanced Trends in Computer Science and Engineering 9(4). DOI:10.30534/ijatcse/2020/351942020. 2020.