Stephen Crawford

Tuguldur Davaadorj

Financial Machine Learning

Introduction

The weak form of the Efficient Market Hypothesis pioneered by Burton Malkeil (*A Random Walk Down Wall Street*, 1973) posits that all price history of a stock is reflected in its current price. By this logic, it would seem implausible for technical trading strategies to predictably outperform the market. In reality, Malkeil's hypothesis actually comes in three forms. Beyond the weak version, there is also the semi-strong version which suggests that all public information is reflected in a stock's current price, and the strong form which suggests that even non-public information (insider information) is reflected in a stock's current price. At first glance, it should be apparent that the semi-strong and especially the strong form of Malkeil's Efficient Market Hypothesis are effectively disproven by the practices and results of large investors such as Warren Buffet and the regulations of the U.S. Securities and Exchange Commission. Unaddressed and lingering however, is the possibility that the weak form of the Efficient Market Hypothesis remains true. That is, it may very well be the case that technical trading strategies are unable to beat the market…

The modern economics student may then wonder: "If technical trading is ineffective, then why do so many high-powered funds practice it?" As pondered in this hypothetical, the findings of research such as that conducted by Smith et al. (*Sentiment and the Effectiveness of Technical Analysis: Evidence from the Hedge Fund Industry*, 2016) would seem counterintuitive to the view argued in Malkeil's seminal literature. Likewise, the immense quantity of established technical indicators published on commonly used resources such as Investopedia, would suggest that there is a very real, if somewhat undulating, view that technical analysis is fruitful.

Returning to the frame of the modern economics student, consider your concerns on the efficacy of technical trading now assuaged. You feel that until a preponderance of evidence is provided to support the weak form of the Efficient Market Hypothesis, you would be foolish to forsake what appears to be a potentially profitable venture which requires little more than a computer and a brokerage account. You may still wonder however, with so many technical indicators being used, which are the most suitable? Is there a best indicator? If not, are there some combinations that work together better than alternatives? Fortunately, this paper discusses a formulaic comparison and analysis of seven different technical indicators with the intention of determining which combinations are the most profitable. Specifically, this paper compares the results of stock trading models trained to use: Bollinger Band Percentage; On-Balance Volume; RSI; Williams

Percent Range; Aroon Oscillator (Up and Down); and Stochastic Oscillator to make trading decisions.

This research was conducted in order to further expand the authors' own understanding of technical indicator efficacy as well as to provide an intuition into which indicators function the best in tandem. Similar studies have been conducted with the goal of analyzing the incorporation of technical analysis strategies into preexisting fundamental analysis approaches. Examples of this previous work include de Souza et al. (*Examination of the profitability of technical analysis based on moving average strategies in BRICS*, 2018). Other works have also looked at the role of technical indicators in price forecasting foreign stock markets. Such work is presented by Lam et al. (*Value Premium and Technical Analysis: Evidence From the China Stock Market*, 2019). Finally, other publications have verified that complex technical indicators are more effective in training machine learning models to develop trading strategies than pure price statistics: McHugh et al. (*Technical indicators for energy market trading*, 2021). In each of these distinct categories of related research, the results suggest that technical indicators offered some form of benefit to the trader. The work presented in this paper is then important for the further insight it provides into both the overall efficacy of technical trading–as a further step towards countering the weak form of the Efficient Market Hypothesis–and the role of different technical indicators in training machine learning models to trade on the stock market.

Methods

Experimental Model

This research makes primary use of a Deep Q Network for the purpose of generating training strategies. The Deep Q Network operates in a custom tailored stock environment which provides all possible combinations of the indicators. A Deep Q Network was selected as the primary learning model for this research because of its ability to accept continuous state features. This means that the stock environment can simply provide the Deep Q Network real indicator values as opposed to quantized, discrete state values as a traditional Tabular Q-Learner would require. Likewise, Deep Q Learning has the benefit of allowing the learner to determine the best actions for itself as opposed to other machine learning methods which require *a priori* knowledge of the correct outcome. An example of this latter group of machine learning algorithms would be any sort of decision tree or K-Nearest Neighbor algorithm. As instance-based learners they require a concrete outcome to direct their decisions.

Experiment Implementation

The specifics of the particular Deep Q Network used for this research are a combination of experience replay and Double Deep Q Learning. The implementation of the network used for this research borrows inspiration from both of these techniques but divorces some aspects in favor of a simpler implementation framework. As with experience replay, the learner makes use of a fixed-length memory which it fills with training tuples as they occur. Specifically, the fixed length memory is able to hold 100,000 experience tuples of (state, action, next state, reward). This memory is then sampled from as with true experience replay; however, the sampling is not done to produce mini-batches. Instead, a single experience is simply sampled from the memory each time.

As with Double Deep Q Learning, an actor network and a target network are used to facilitate convergence of the actor network to the correct optima. However, instead of the newer, completely decoupled version of Double Deep Q Learning, this implementation simply makes use of the target network as a slower periodic copy of the actor network.

Given the suitability of the problem to Deep Q Learning, there were few other machine learning algorithms which were considered. The two that were include advantage actor-critic as well as *asynchronous* advantage actor-critic which was considered in light of the computational load of this research. Both of these alternative Deep Q Learning frameworks were decided against due to the time constraints of the research period and their relative complexity. For both instances, it would be necessary to be able to create a utility function which could then be used by the critic network to estimate state-utilities. In the case of the stock environment, where the future value of a position is relatively unknown, there was no obvious utility function to be used.

Beyond the selection of a model architecture there were several specific considerations which were made to best accommodate the analysis of seven different technical indicators. One of the most obvious problems resulting from the use of Deep Q Learning is the computational and time cost presented in the training. Given seven indicators, the process of training a model to use each combination quickly becomes intractable. Originally, the intention was to compare the results of every combination of the indicators. This means experiments would be run for every indicator by itself, then every pair of indicators, then every group of three, etc. The resulting number of combinations is a total of 127 learners which would have to be trained. Using 500 trips to train each learner as the authors had done in previous works, would result in 63,500 total trips through the date range. At an average run-time of 20 seconds, this would require 14.7 days to run. This obviously was not feasible so a number of accommodations were made in order to shorten the runtime.

First, the number of trips used for training was cut from the previously used 500 trips to 200 trips. This scaled the total runtime down to 40% of its original value. This number was selected as it was believed to be a large enough number for the Deep Q Learners to still train a significant portion. After all, the intention of this research is not to train the best specific instance of a learner but instead compare the impacts of the indicators. The second most important change made was the use of manually implementing parallel computing. Because the time to run would still be far too large, the largest jobs–i.e. the learners being trained on 7C2, 7C3, 7C4, 7C5–were split in half by creating two instances of scripts which would train the learners and having one learner train on approximately half of the combinations and another train on the other half. This was done in a manner that allowed the results to then be recombined as CSVs. Finally, the different learner instances were run as four separate parallel computing scripts on Bowdoin College's Hyperformance Computing Grid. Each learner was then able to make use of a NVIDIA 3080 graphics card and CUDA tensorflow integration. After these accommodations were made, the jobs were then queued on the grid according to the researchers' own opinion of interest. That is, the jobs were queued so that the 7C3 and 7C4 jobs would run first followed by the lesser combinations.

Dataset

This section describes the data used for experimentation. The data used to study the efficacy of different technical indicators was sampled from a dataset of the stocks in the Russell 1000 Index. Specifically, only Disney (DIS) was looked at over a four year date-range (2018-2022). This data was selected due to the researcher's familiarity with the data from previous work and the purpose of the research. Because the research was focused on the efficacy of the indicators compared to one another, the specific stock data selected would appear irrelevant so long as the stock was not incredibly anomalous. Disney's stock price shows reasonable behavior over the time period in question and so therefore made a strong candidate for use. The data was not preprocessed in any way, it was simply extracted from the original comma-separated values file which it was contained in. Other candidate datasets which were considered were the SPY ETF over a similar date range or Disney over a time period which spanned the 2008 Recession (2008-2012). Ultimately, both of these alternatives showed no real benefit over the originally selected Disney stock and date range.

Discussion

Experimentation

In order to determine the most effective technical indicators for training a Deep Q Learner to develop a stock trading strategy we conducted a significant number of trials of the same experimental setup. For each of the different technical indicator combinations, a Deep Q Learner was trained over a date range of 2018-01-01 to 2019-12-31. This date range was run for 200 'trips' with each trip being a run through the training time period. The Deep Q Learner was run using a stock environment which would provide experience tuples as aforementioned. These experience tuples were used by the learner to calculate a loss function based on the difference between the observations and the predicted values. The loss function used was Huber loss with the target network providing the 'correct estimate' portion of the traditional Q Learning loss function and the actor network receiving all the updates. Huber loss has the property of being quadratic for smaller values and linear for larger values. Huber loss was selected based on the documentation of Tensorflow and Keras which made use of the loss function when calculating the loss for the Gym CartPole problem. When the loss was received by the learner, gradient descent was used to adjust the Q network weights to minimize the loss. The Deep Q Network was implemented using Keras and consisted of seven input neurons, leading to a densely connected layer with half as many neurons as the input layer. Finally the output layer of the network was simply the three possible actions the learner could take: buy, hold, or short. The shallow network was selected given the relatively simplicity of the problem.

Beyond the specifics of the training process, the experimentation involved in the research was concerned with the results of the different learners compared to one another. The primary goal being determining which technical indicator was most effective in training a Deep Q Learner to develop a stock trading strategy. For each of the 127 possible combinations, the learner was trained on the in-sample date range and then tested both in-sample and out-of-sample. Testing both in-sample and out-of-sample allowed for a more sophisticated picture of each learners' behavior. If only in-sample testing was done, it would be unclear whether the Deep Q Learner had simply overfitted to the specific time period. Because the use-case is stock trading, a model which can only perform well on data it has already seen is useless.

In order to analyze and interpret the different combinations efficacy, the program was set up to create CSVs each containing the test results for the corresponding indicator counts. For example, a CSV might be titled: *Three_Indicator_Results_1_in_sample*. This nomenclature denotes that the CSV corresponded to the three indicator combinations, was the first of the parallel computing processes to complete, and contained the model's in-sample test results. The test results took the form of the model's portfolio value over the test range. Portfolio value was calculated as: $cash\ +\ shares * share\ price$. The initial starting cash was $200,000. Transaction fees were not enforced as the intention was to simplify the problem as much as possible in order to determine the extent to which the indicators influenced the performance. After running the experiments, the CSVs were recombined and further analyzed to look for the best performing indicator combinations. Specifically, the data for each of the indicator counts was examined for the best performing indicators as of the end of the test periods (which indicators made the most money) as well as trends in the better performing indicators across combination sizes.

Prior to analyzing the results, it was expected that learners trained on a greater number of indicators would show, on average, better performance both in and out-of-sample. It was also expected that for smaller combinations of indicators, the best performing combinations would contain On-Balance Volume, Bollinger Bands, and one of the Oscillators. The rationale behind this expectation was that these three indicators would represent the maximum amount of information that the learner could garner from only three specific indicators. On-Balance Volume is the only indicator which interprets volume data, Bollinger Bands offers logically-sound measures of the stock's price, and Oscillators generally interpret market swings.

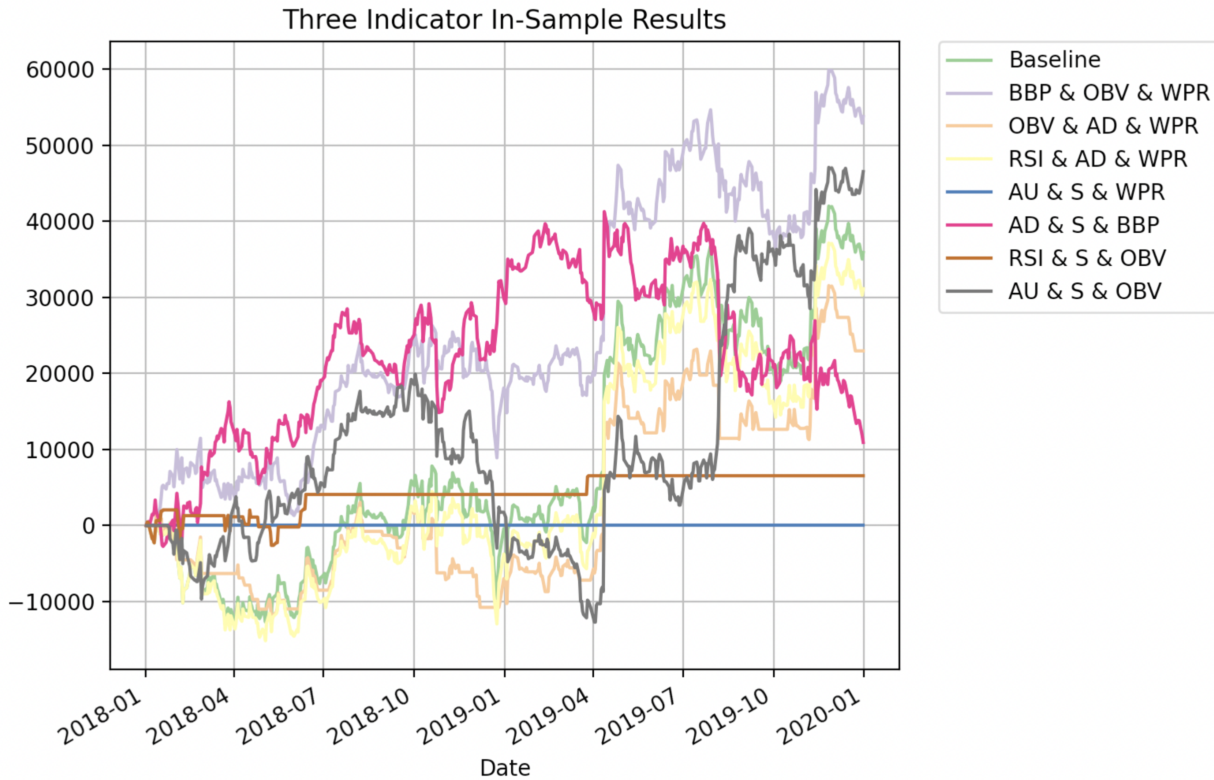
Results

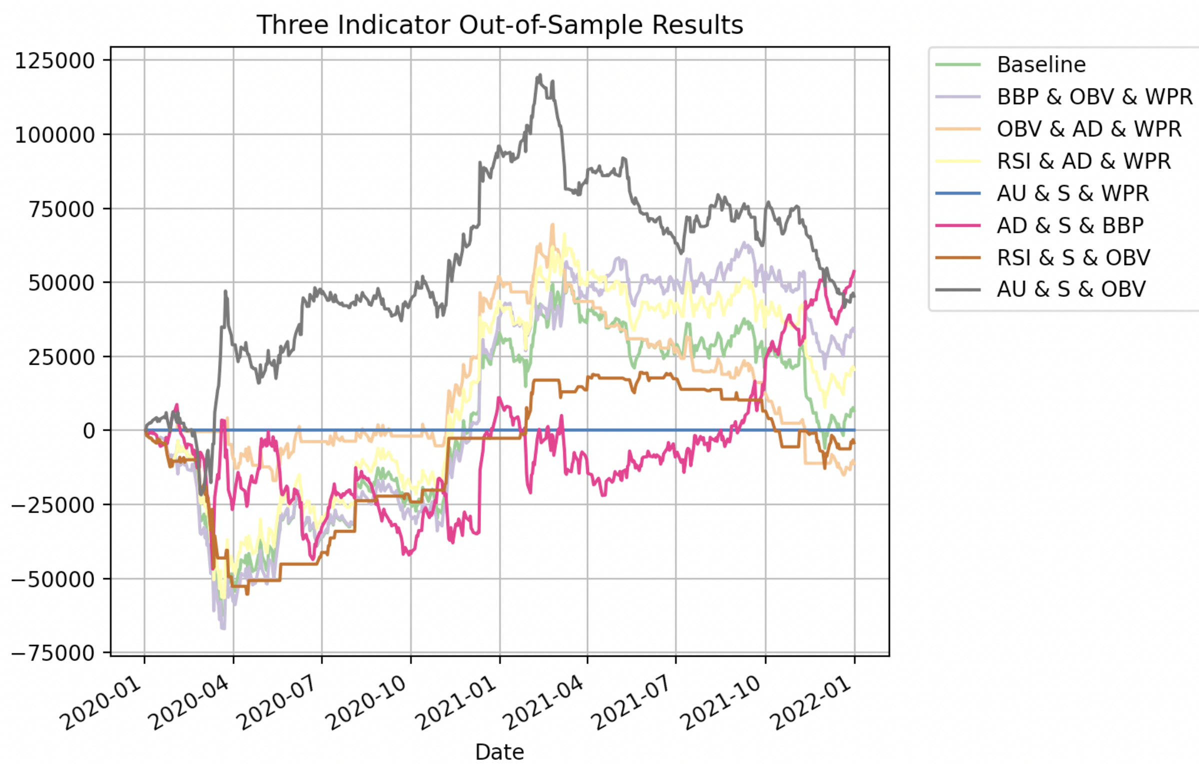
Two types of conclusions can be drawn from this research: those concerning indicator efficacy and those concerning the machine learning methods used. The following two subsections discuss each of these categories respectively.

Indicator Efficacy

As mentioned, the fundamental goal of this research is to determine the most effective indicator combinations. With regard to the experimental setup, this means that the goal is to determine which learner (each being trained on a different combination) is most effective. Before discussing the specific results it should be reiterated that the *a priori* assumption of this research was that there would be an indicator(s) showed consistent performance. A set of indicators that perform well would be expected to consist of subsets that also performed well. Finally, before discussing the specific findings for each combination size, some general notes: due to time constraints, the training for combinations of size 1 and 2 were not performed; after running for 200 trips, a majority of the indicator combinations "converged" to the baseline. These converged combinations would hold a long position on day one and hold, causing it to have the same rewards and overall as the baseline. Therefore, these combinations are not included in the following plots due to redundancy. While it is interesting to see how frequently the DQN developed a trading strategy which amounted to "buy and hold," it is ineffective in answering the core question of this research. Finally, for each combination, abbreviations were used for the indicators where "WPR" is Williams Percentage Rate, "BBP" is Bollinger Bands Percentage, "OBV" is On-Balance Volume, "AU" is Aroon up, "AD" is Aroon Down and "S" is Stochastic Oscillator.

The following graphs show all the combinations of three, four, five, six and seven indicators that did not equal the baseline and offer analysis of these results.


Three Indicator In-Sample Results
60000
50000
40000
30000
20000
10000
0
−10000
2018-01
2018-04
2018-07
2018-10
2019-01
2019-04
2019-07
2019-10
2020-01
Date
Baseline
BBP & OBV & WPR
OBV & AD & WPR
RSI & AD & WPR
AU & S & WPR
AD & S & BBP
RSI & S & OBV
AU & S & OBV

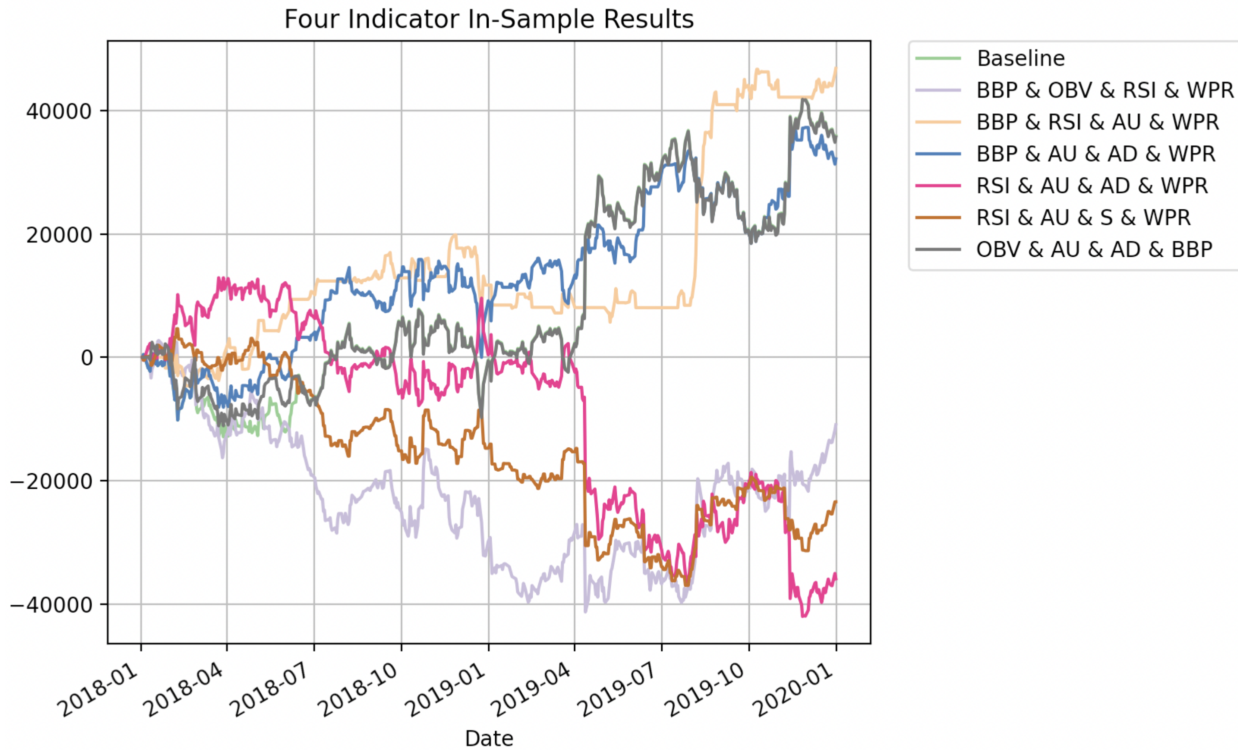

Three Indicator Out-of-Sample Results
125000
100000
75000
50000
25000
0
−25000
−50000
−75000
2020-01
2020-04
2020-07
2020-10
2021-01
2021-04
2021-07
2021-10
2022-01
Date
Baseline
BBP & OBV & WPR
OBV & AD & WPR
RSI & AD & WPR
AU & S & WPR
AD & S & BBP
RSI & S & OBV
AU & S & OBV

These plots represent the in-sample and out-of-sample indicator combinations which did not converge to the baseline. As can be seen a few notable observations can be made. Based on the graphs above, in order of most profit, the combinations of (AD, S, BBP), (AU, S, OBV), (BBP, OBV, WPR) and (RSI, AD, WPR) outperformed the out-of-sample baseline (this can be further verified by checking the numeric value in the appropriate CSV files). Whereas, for the in-sample data, (BBP, OBV, WPR) and (AU, S, OBV) outperform the baseline. From this result, there is no correlation for specific indicators that perform the best. While OBV, AD, BBP, WPR and S are part of multiple of the out-of-sample combinations which performed the best, it is not evident that this is due strictly to the indicators. For instance, (OBV, AD, WPR) underperforms for both the in-sample and out-of-sample data. This lack of correlation in indicators is further evident when looking at the number of converged combinations per indicator, shown in the table below.

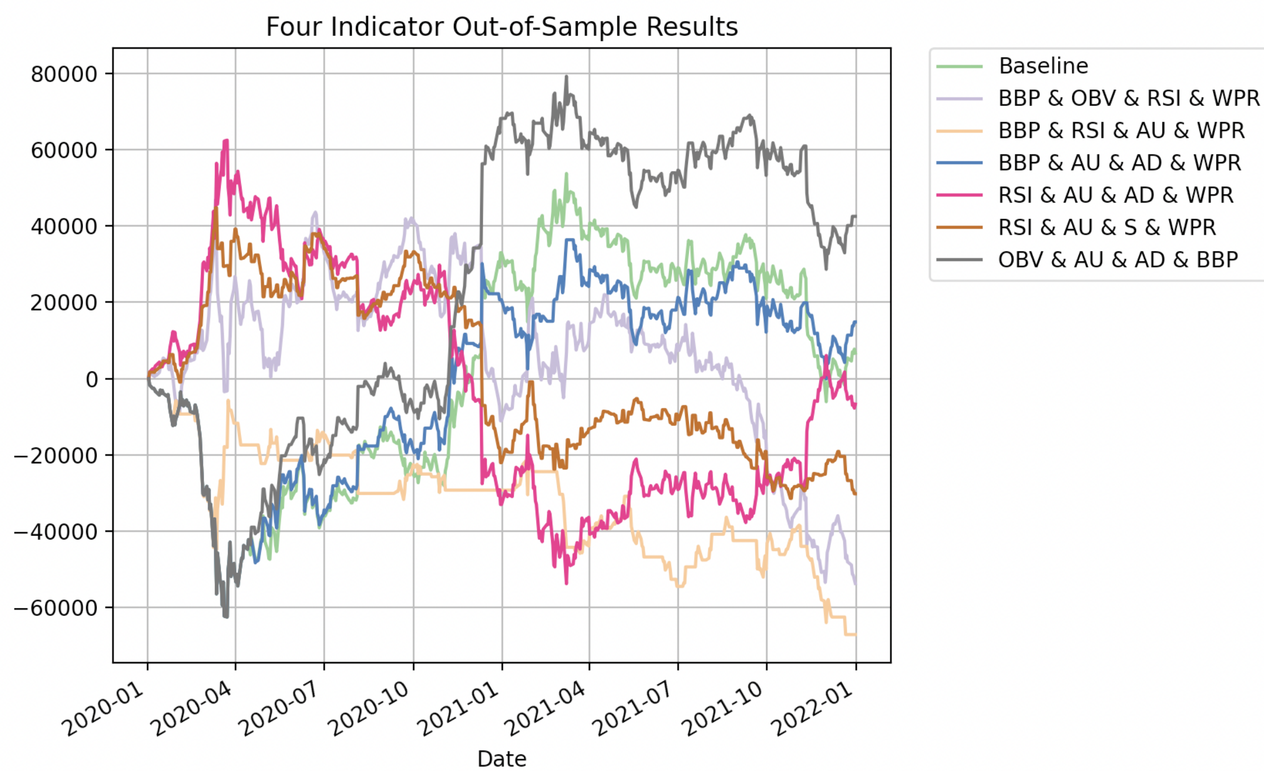
Indicator	Number of Converged Combinations of size 3 containing the indicator (Max 15)
Williams Percentage Rate	11
Bollinger Bands Percentage	13
On-Balance Volume	11
RSI	13
Aroon Up	13
Aroon Down	12
Stochastic Oscillator	11

With at least 11 out of 15 of each indicator's combination converging, it is impractical to compare the few combinations that did work and determine the best indicator and best combination. It appears that WPR, OBV, AD and S converged the least, while BBP, RSI, AU converged the most. In order to further develop an understanding of the role of specific indicators, it is convenient to analyze the larger-size combinations while keeping these findings in mind. If larger sets containing (AD, S, BBP), (AU, S, OBV), (BBP, OBV, WPR) and (RSI, AD, WPR) out-perform the market, it is reasonable to conclude these indicators work well in tandem. Likewise, if combinations containing WPR, OBV, AD and S are also outperforming the market for combinations of four, five, six and seven, then we can conclude that WPR, OBV, AD and S might be an estimable indicator(s).

With this in mind, we look at the combinations of four that did not converge to the baseline for both the in-sample and out-of-sample data:



Four Indicator In-Sample Results
Baseline
BBP & OBV & RSI & WPR
BBP & RSI & AU & WPR
BBP & AU & AD & WPR
RSI & AU & AD & WPR
RSI & AU & S & WPR
OBV & AU & AD & BBP
40000
20000
0
−20000
−40000
2018-01
2018-04
2018-07
2018-10
2019-01
2019-04
2019-07
2019-10
2020-01
Date



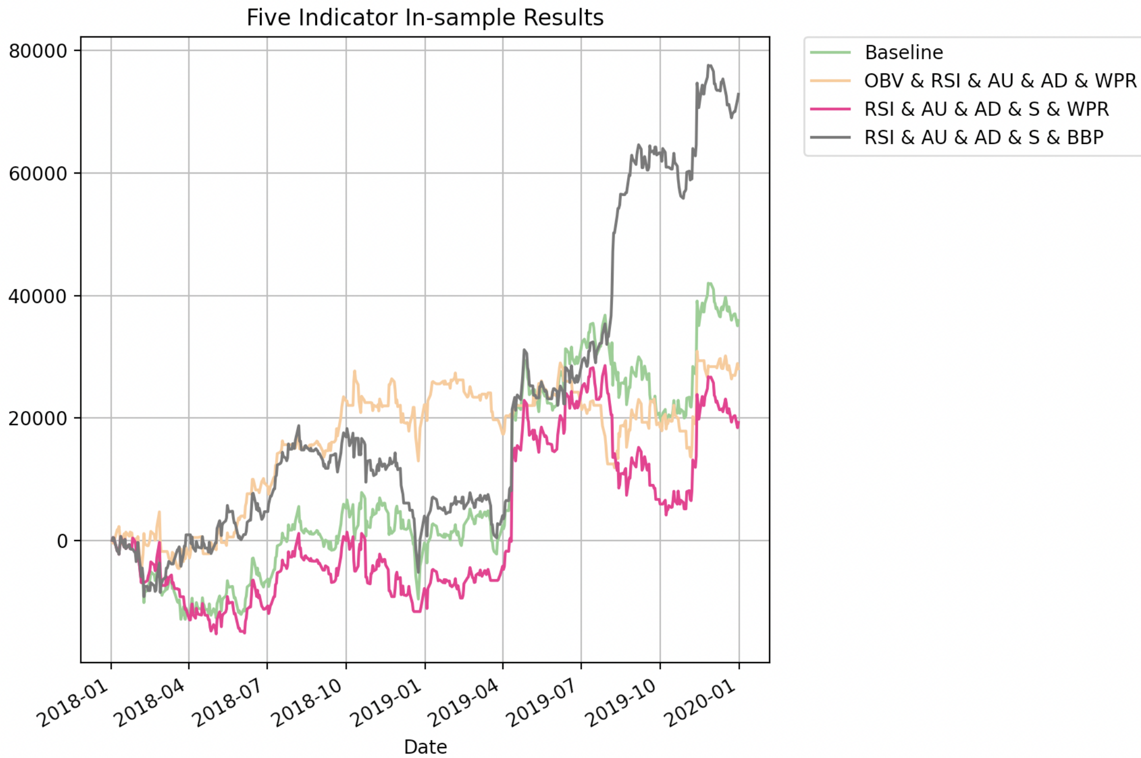
Four Indicator Out-of-Sample Results
Baseline
BBP & OBV & RSI & WPR
BBP & RSI & AU & WPR
BBP & AU & AD & WPR
RSI & AU & AD & WPR
RSI & AU & S & WPR
OBV & AU & AD & BBP
80000
60000
40000
20000
0
−20000
−40000
−60000
2020-01
2020-04
2020-07
2020-10
2021-01
2021-04
2021-07
2021-10
2022-01
Date

In the four-indicator combinations shown above, the combinations that outperformed the out-of-sample data were (OBV, AU, AD, BBP) and (BBP, AU, AD, WPR). For in-sample, the baseline is obscured by the overlapping (OBV, AU, AD, BBP) line and only (BBP, RSI, AU, WPR) surpasses it. This results conflicts with our previous finding for combinations of size 3. The outperforming three-indicator combinations were not a subset of the outperforming buckets of 4, indicating a lack of relation between the subsets of indicator sets. For combinations of 4, the AU, AD, BBP indicators are included in both of the out-of-sample winners, where AD and BBP both performed well for combinations of 3. On a lesser note, OBV and WPR also are part of the combination that perform well. Combinations containing AD, BBP, AU, WPR and OBV should perform well, based on the results. However, this is not definite since, once again, there is a high amount of converged combinations, as shown in the table below.

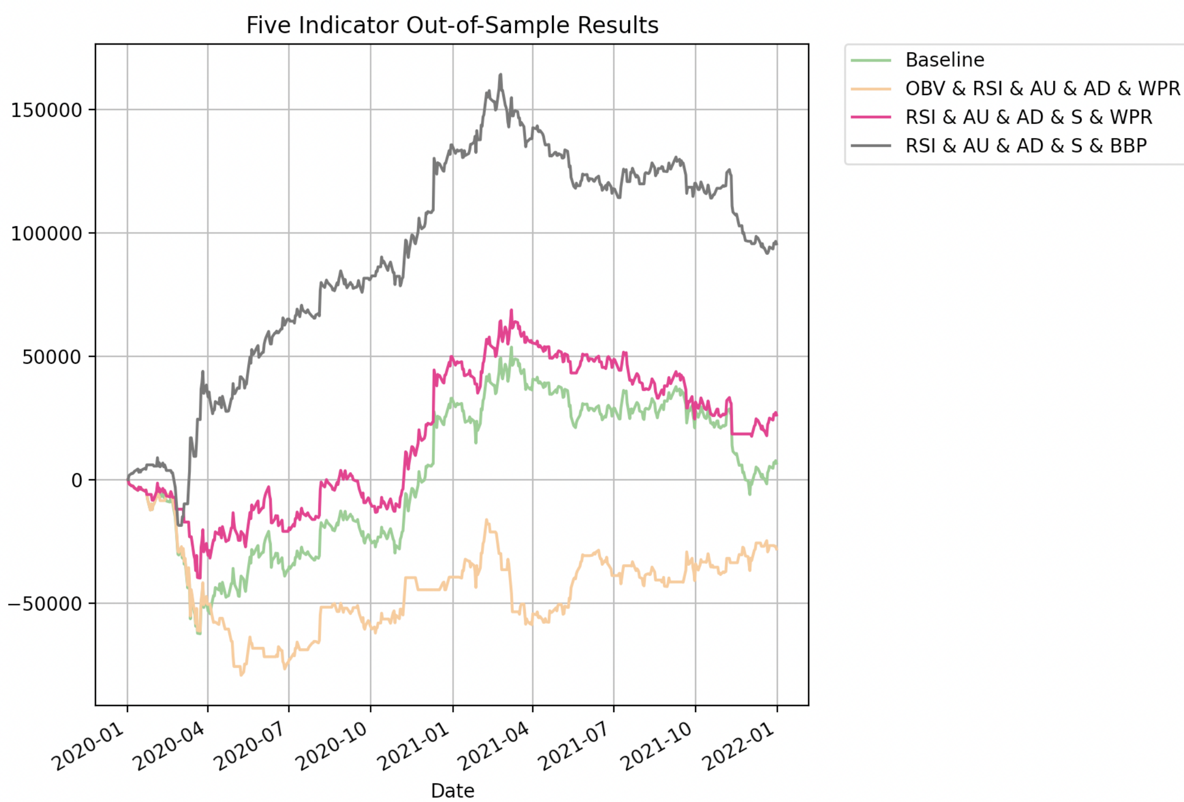
Indicator	Number of Converged Combinations of size 4 containing the indicator (Max 20)
Williams Percentage Rate	15
Bollinger Bands Percentage	16
On-Balance Volume	18
RSI	16
Aroon Up	15
Aroon Down	17
Stochastic Oscillator	19

At Least 15 out of 20 of each indicators' combination converge to the baseline. WPR once again converges the least. However, S and OBV converge among the most, meaning that there is no clear pattern. While it appears that some combination of OBV, AU, AD and WPR work the best, there are no evident relations between the combination of 3 and 4 results due to the convergence. Once again, it is best to avoid broad conclusions given the results somewhat confounded nature. Turning attention to the combinations of 5 indicators only further obscures any underlying trends.

The graphs for 5 combinations that did not converge for both in-sample and out-of-sample are included below, where only 3 out of 21 possible combinations did not converge.



Five Indicator In-sample Results
80000
60000
40000
20000
0
2018-01
2018-04
2018-07
2018-10
2019-01
2019-04
2019-07
2019-10
2020-01
Date
Baseline
OBV & RSI & AU & AD & WPR
RSI & AU & AD & S & WPR
RSI & AU & AD & S & BBP



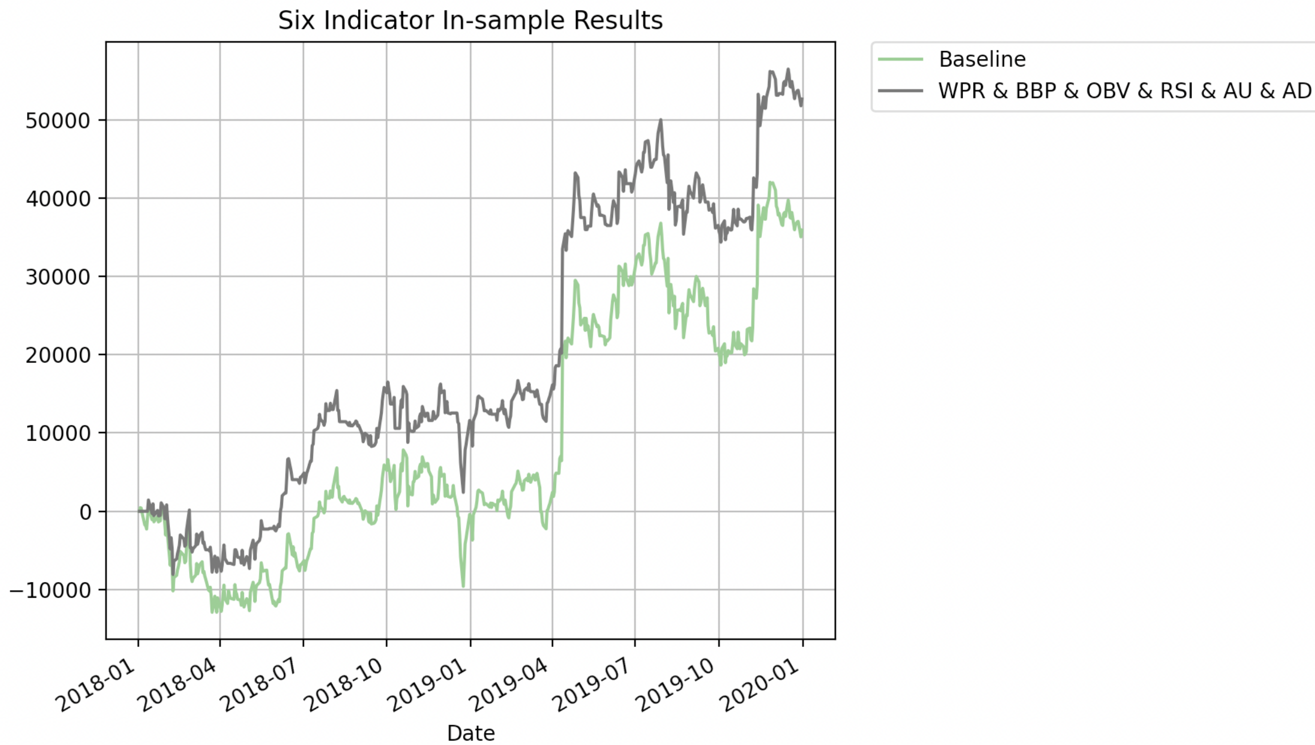
Five Indicator Out-of-Sample Results
150000
100000
50000
0
−50000
2020-01
2020-04
2020-07
2020-10
2021-01
2021-04
2021-07
2021-10
2022-01
Date
Baseline
OBV & RSI & AU & AD & WPR
RSI & AU & AD & S & WPR
RSI & AU & AD & S & BBP

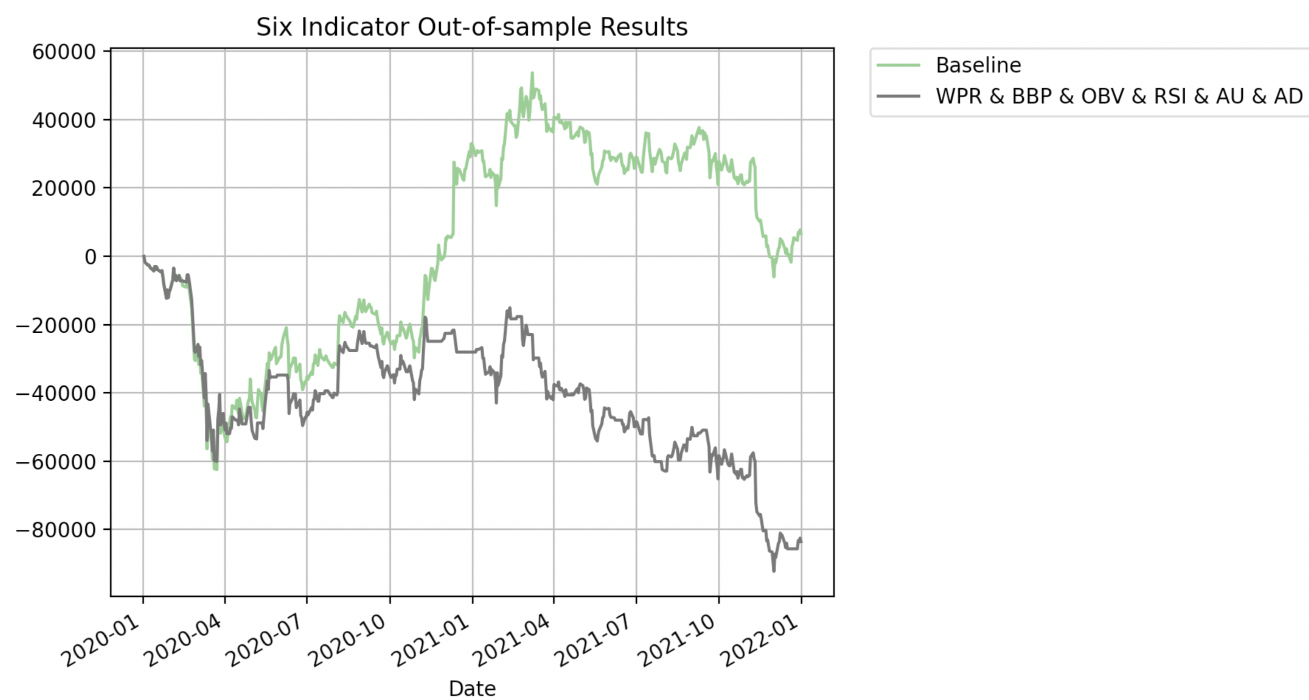
For the in-sample and out-of-sample five-indicator combination diagrams shown above, the combinations of (RSI, AU, AD, S, BBP) and (RSI, AU, AD, S, WPR) outperformed the out-of-sample baseline and only (RSI, AU, AD, S, BBP) outperformed the in-sample. AU, AD and RSI are the common outperforming indicators. AU and AD were part of the subsets that performed well for the 4 indicator sets. OBV, which performed well for 3 indicator sets and 4 indicator sets shows far worse performance as a member of the five-indicator combinations. WPR and BBP perform well once again. However, there is a high degree of uncertainty because a majority of the combinations converge to the baseline, rendering their results inconclusive.

Indicator	Number of Converged Combinations of size 5 containing the indicator (Max 15)
Williams Percentage Rate	13
Bollinger Bands Percentage	14
On-Balance Volume	14
RSI	12
Aroon Up	12
Aroon Down	12
Stochastic Oscillator	13

When comparing the table above to the combinations of 3 and 4, there is no evidence of a particular indicator that causes convergence. The range of converged combinations is low. This likely indicates that the convergence is more due to the environment and Q-trader than the indicators. This pattern continues for the combinations of 6 and 7 results.

The results for 7 indicators is one combination of all 7 indicators that converged to the baseline (No graph is provided, since it would be oscucurred by the baseline for both the in-sample and out-of-sample). The graphs for the combination of 6 indicators that did not converge to the baseline are provided below. Out of the 7 combinations, only 1 combination did not converge.





The combination of (WPR, BBP, OBV, RSI, AU, AD) outperforms the in-sample baseline but considerably underperforms the out-of-sample baseline. From prior combination size results, S did not have a clear pattern of the most convergence, but it was the only indicator without a converging combination. This combination contains all the prior indicators (AD, AU, WPR, BBP) that regularly outperformed the baseline. The results indicate that the combination of the "good" indicators does not necessarily result in a better result. Additionally, we hypothesized that a greater basket of indicators would return greater profit. This is disproven by this result since no combination of 6 indicators outperformed the out-of-sample baseline while there are 3, 4, 5 indicator combinations that beat it out. The following table provides all the combinations that outperformed the out-of-sample baseline and the reward to compare the best combinations.

Indicators	Total Reward Amount
Baseline	$6,690.00
(AD, S, BBP)	$53,650.00
(AU, S, OBV)	$45,340.00
(BBP, OBV, WPR)	$33,490.00
(RSI, AD, WPR)	$20,530.00
(OBV, AU, AD, BBP)	$42,500.00
(BBP, AU, AD, WPR)	$14,840.00
(RSI, AU, AD, S, WPR)	$26,220.00
(RSI, AU, AD, S, BBP)	$95,550.00

From the table above, (RSI, AU, AD, S, BBP) performed the best. Considering that none of the subsets of this combination outperform the baseline, this result might be due luck more so than a precisely fine-tuned set of an indicator basket. Not to mention, this combination contains RSI and S, which generally did not perform well. To map the efficacy of each individual indicator, the following table shows the amount of non-converged combinations across all combinations of size 3, 4, 5, and 6.

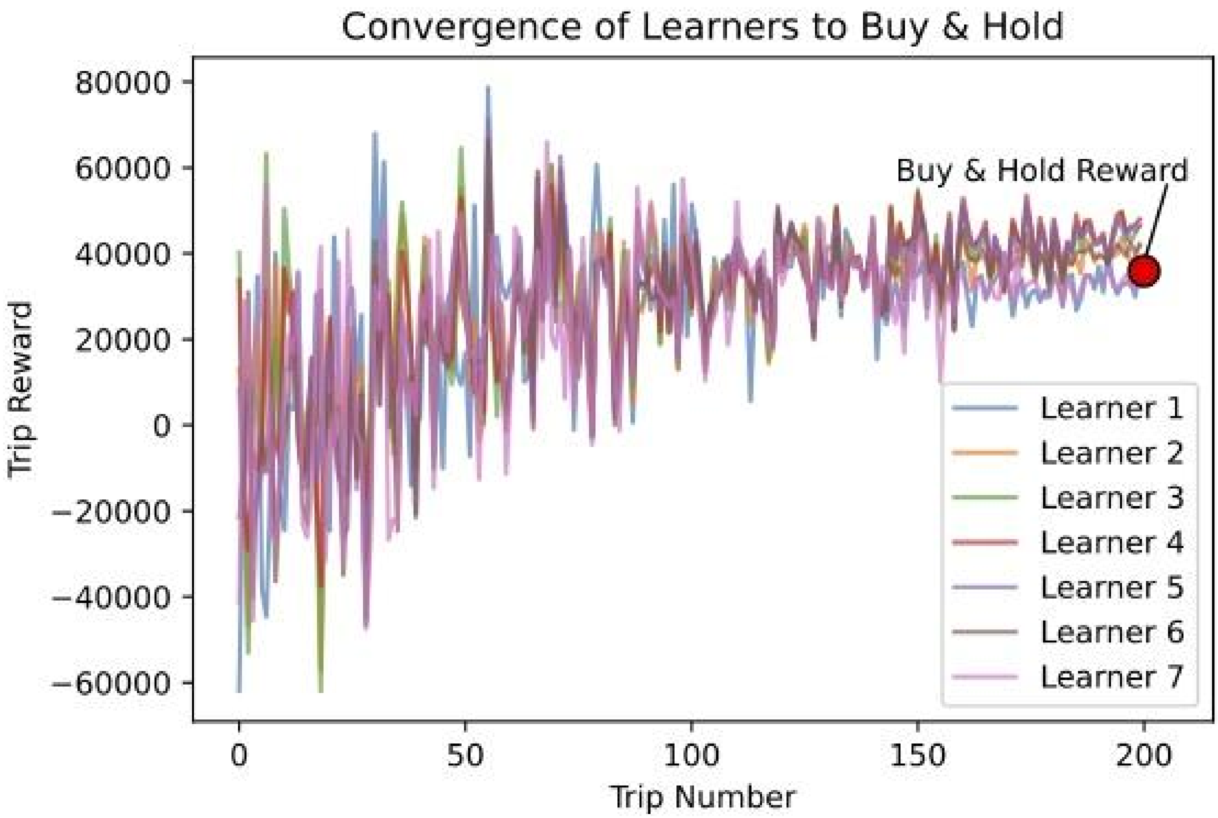
Indicator	Number of Non-converged Combinations containing the indicator (Max 20)
Williams Percentage Rate	4
Bollinger Bands Percentage	5
On-Balance Volume	3

RSI	3
Aroon Up	5
Aroon Down	5
Stochastic Oscillator	3

The best performing indicators in terms of non-convergence were AU, AD, and BBP, which are all elements of the best performing subset. Aside from obvious shortcomings such as a constricted date range, trading only a single stock, using only 7 indicators explored (10+ indicators often used in hedge funds), the biggest issue is that a majority of the trained combinations converge to the baseline. This reduces the amount of comparisons we can make to determine the best indicator since we cannot assume the efficacy of a converged set of indicators. Out of the 99 possible combinations we looked at, only 17 of the combinations did not converge to the baseline. In the next section, we look at possible reasons for this convergence and the overall efficacy of the learner.

Learner Analysis

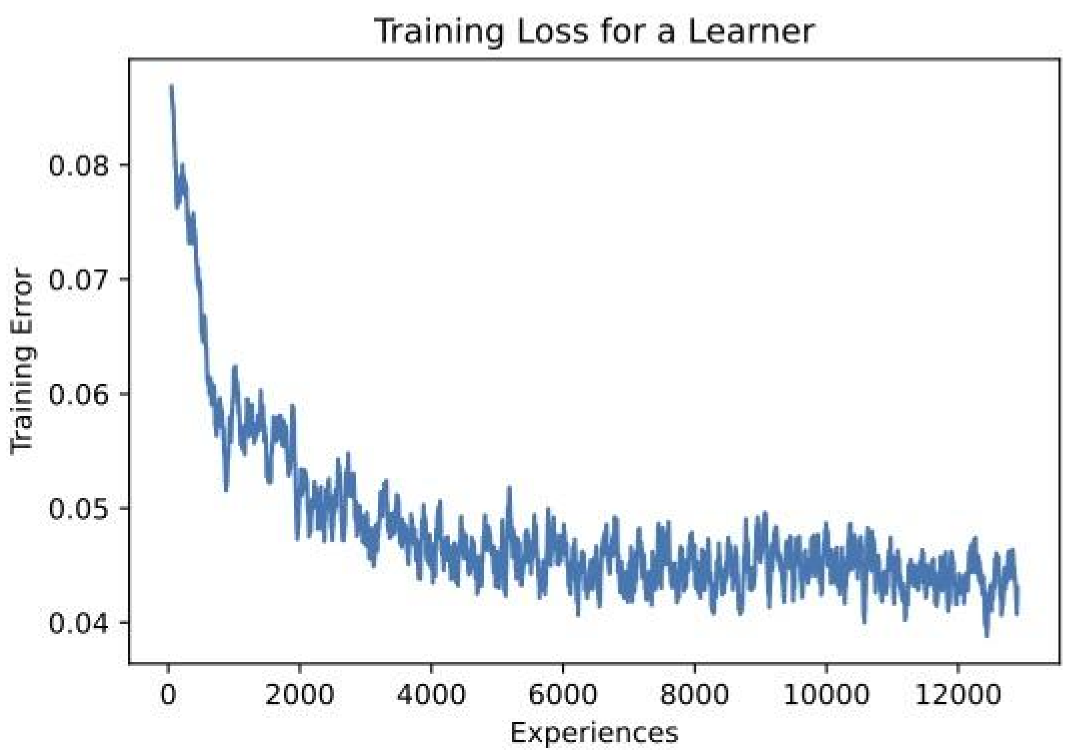
Just as the results provided an intuition into the efficacy of different technical indicator combinations, they also elucidated a better understanding of learning architectures. Two notable trends complicate the overall process of using a DQN to train a stock trading model. Specifically, many of the learner instances (many different combinations of indicators) saw optimal policies which converged by 200 trips to approximate a buy and hold strategy.



This diagram shows various different learners trained with different indicator combinations which all generally trend towards a buy and hold strategy as the number of trips trained increases.

The convergence of many different learners to a buy and hold strategy was not an expected outcome and was originally thought to be the result of faulty learners. DQNs are regularly used to learn policies for much more complicated tasks, and when transaction costs were set to \$0 (fixed and floating) the learners were expected to develop rapid trading strategies. Instead, the learners frequently converged to pursue more consistently favorable outcomes instead of enacting riskier policies. Though the learners were not directly punished for making trades, examination of the training trip rewards shows that even learners which had experienced *greater* rewards for a single trip, generally converged to policies which had less potential for gain but also less risk. This indicates that in order for a DQN to be used to develop a high-frequency trading policy, it is important to motivate the learner to make trades simply for their own sake. In the case of this research, it is shown that the absence of negative reinforcement does not equate to the presence of positive reinforcement. Plainly, just because the learners did not have to pay transaction fees, did not mean that they would rather make trades and risk losing money when they could just buy and hold for the entire training period.

Further support for this theory is shown when examining the loss during training of a single learner. As can be seen in the below diagram, the loss of the learner quickly reaches a horizontal asymptote. This not only shows that the learner does *learn*, but also that it begins to converge to an 'optimal' policy relatively quickly. Recall, as defined in this research, the loss function of the learner is the difference between the actor network's predicted Q-value and the reward plus the weighted target network's Q-value. That the loss quickly decreases to an apparent asymptote indicates that this difference begins to converge around the 10th trip (there are 503 experiences per trip).



This diagram depicts the training loss of a learner as it undergoes training trips. Notice the horizontal asymptote approached around 5000 experiences.

Ultimately, it seems obvious that in order for a DQN learner to develop a trading strategy, it is important that techniques are used to delay convergence and motivate the learner to make frequent trades.

Conclusion

Following the reasoning outlined in the Results section, several notable conclusions can be made. For technical indicators, due to the high number of convergence, we cannot accurately claim a specific set of indicators to be the best. Based on our results, we saw that (RSI, AU, AD, S, BBP) performed the best, where combinations with AU, AD, and BBP had the highest amount of combinations that outperformed the out-of-sample baseline. Experimenting in a similar manner with other learners or a high-frequency trader could result in less converged outcomes that would increase the amount of comparisons available to accurately estimate the impact of the subsets for a given combination of indicators. In turn, this would allow for regression analysis of the result, which could accurately quantify the impact of each indicator. Another future line of research could be comparing the characteristics of indicators such as price, volatility, momentum etc. to see which (combination of) characteristics perform the best.

With regard to the methodology, as mentioned, it seems important that DQNs used to determine high-frequency trading strategies be trained using a reinforcement structure that encourages trading. Without this, it is likely that the optimal policy determined by the learners, regardless of which technical indicators are used, will be one of buy and hold. In light of this, it would be fruitful for future research to be conducted into the efficacy of technical indicators when trading is enforced. That is, which indicators would perform best if the learner must make some trade every day? Other interesting options for future research are a similar comparison of technical indicator combinations using alternative machine learning learners or further modification of a DQN so that it could better rationalize about the opportunities it *loses* through holding.

Citations

Catherine McHugh, Sonya Coleman, Dermot Kerr, Technical indicators for energy market trading, *Machine Learning with Applications*, 6, 2021, https://doi.org/10.1016/j.mlwa.2021.100182.

de Souza, M.J.S., Ramos, D.G.F., Pena, M.G. *et al.* Examination of the profitability of technical analysis based on moving average strategies in BRICS. *Financ Innov* 4, 3 (2018). https://doi.org/10.1186/s40854-018-0087-z

Lam KSK, Dong L, Yu B. Value Premium and Technical Analysis: Evidence from the China Stock Market. *Economies*. 2019; 7(3):92. https://doi.org/10.3390/economies7030092

Malkiel, Burton Gordon. *(2003). A random walk down Wall Street : the time-tested strategy for successful investing.* New York :W.W. Norton

Smith, D., Wang, N., Wang, Y., & Zychowicz, E. (2016). Sentiment and the Effectiveness of Technical Analysis: Evidence from the Hedge Fund Industry. *Journal of Financial and Quantitative Analysis, 51*(6), 1991-2013. doi:10.1017/S0022109016000843