The Integration of Technical Indicators into a Deep Q-Learning System to Increase Profitability in Simulated Stock Market Trading

Matthew Kahn

Somers High School

Table of Contents

Acknowledgements

I would like to thank my Science Research teachers, Mr. Maelia and Mrs. Lepkowski, my computer science teacher, Mr. Soni, as well as my family and peers, for their constant support and for guiding me through my research.

Abstract

Although stock market trading has been successfully conducted in domestic and international markets using deep Q-learning, systems fusing deep Q-learning with technical indicators have been largely absent from the literature. Considering the innate difficulty of stock market trading, as well as the sheer number of individuals invested in the stock market, the full exploitation of all analytical tools is essential for researchers and investors alike. Separately, technical indicators and machine learning systems are widely used as they can process large, convoluted datasets, and make them far easier to interpret. In this paper, a unique deep Q-learning system is described, which substitutes the typical long- and short-term reward system for a more complex technical-indicator-based hybrid system. This updated system automatically weighs technical indications and simulates stock market trading in order to develop unique investment strategies. The results indicate that the updated reward system generates significantly higher returns relative to the traditional method and the benchmark buy-and-hold method. Additionally, the final state of the genetic algorithm indicates that the Money Flow Index (MFI) is the tested indicator most indicative of future price changes.

Tables & Figures

Introduction

According to Forbes (2020), nearly 55% of Americans are invested in the stock market. Undoubtedly, all of these investors wish that they could predict the stock market. Unfortunately for these investors, this is not the case as forecasting the stock market has remained one of the most challenging endeavors in financial analysis. The unpredictability of the market can largely be attributed to the numerous, complex factors that affect the behavior of a stock, namely politics, investor sentiment, and broader economic conditions (Nazário et al., 2017). Core economic theories exist because of the chaotic nature of the stock market, namely the Efficient Market Hypothesis (EMH) and Random Walk Theory, which conclude that excessive returns cannot be earned using exclusively time series data. Despite these conclusions, the use of technical indicators for conditioning trading strategies has granted investors significantly better investment results than the benchmark buy-and-hold method (Wang et al., 2007). Additionally, advancements made in machine learning, paired with readily available stock market information, have allowed researchers to create high-performing trading models (Wang et al., 2017; Li et al., 2019).

Technical analysis is a trading methodology used for predicting changes in the market through the investigation of past market data. Technical traders can use individual tools, called technical indicators, to refine chaotic stock market information and more easily interpret market data. In addition to technical indicators, machine learning (ML) can also be used to manage the complexity and volatility of the stock market, allowing traders to more easily find optimal trading strategies (Xiong et al., 2018). A popular and versatile variant of machine learning called deep Q-learning has demonstrably outperformed other variants of machine learning and has been applied to a multitude of tasks, including video game playing (Mnih et al., 2015), autonomous driving (García et al., 2019), and stock market trading (Wang et al., 2017). Deep Q-learning, although present in the literature surrounding stock market trading, has rarely been used alongside technical analytics. Traditional deep Q-learning relies on a formula called the Bellman Equation, the simplicity of which may limit the decision-making ability of deep Q-learning. In this study, a formula based on five popular technical indicators – optimized using genetic algorithms – is substituted in place of the Bellman Equation to identify any potential improvements to deep Q-learning based trading and gauge the relative value of various technical indicators.

Review of Literature

As described by (Zhang et al., 2019), technical analysis (TA) is a fundamental discipline of stock market trading which seeks to understand price trends through the investigation of historical stock market data, including price, volume, float, and volatility. TA revolves around the use of technical indicators, which reconfigure market data for more transparent interpretation. Thanks to the availability of data and the development of open digital resources, technical analysis has emerged as a leading method of market forecasting (Jakpar et al., 2018). Although grandfathered into market analytics, technical indicators have primarily been applied to neural networking only as a substitution for price labels or as supplemental preprocessing (Patel et al., 2015; Dash et al., 2016). As such, the surrounding literature has been largely void of complete hybridization between technical indicators and niche ML variants like deep Q-learning.

As noted by (Wang et al., 2017), machine learning (ML) methods can create profitable trading strategies from chaotic raw data, making such methods extremely useful for analysts. Of the various types of machine learning, a variant known as reinforcement learning (RL) can navigate a complex environment by using a rewards system and a learned decision policy to determine optimal actions or investments (Sah, 2020). RL is particularly useful for the processing of complex systems like the stock market, explaining its recent growth in popularity (Mosavi et al., 2020).

Q-learning, a leading variant of reinforcement learning, determines the highest-quality actions or investments to be taken at any particular state of an environment. Iterating through all possible states and actions, potential modifications to an environment are analyzed to return a reward (which is cumulatively maximized by the agent). Once all combinations of actions are analyzed, an optimal value function, Q(s,a) is then derived and can be used for finding the optimal series of actions to be taken (Watkins & Dayan, 1992). The equation central to Q-learning is known as the Bellman Equation, which was translated to Python for the purposes of this experiment and can be seen below in *Figure 1*.

```
new_Q = current_Q + alpha*((current_reward) + gamma*(max_future_reward))
```

Figure 1. Python Code representing the Bellman Equation (Created by student researcher, 2021).

Regarding the Bellman Equation, learning rate (α) determines the sensitivity of the agent and discount factor (γ) determines the weight of long-term reward. Both learning rate and discount factor are configured manually, and as such are known as hyperparameters. As noted by Wang et al., (2017), the Bellman Equation can be used to approximate the quality of each action at each state. A large underlying problem with textbook Q-learning is that functionality ceases in an unlimited-state environment, such as the stock market. As a solution, researchers supplement Q-learning with deep neural networks (a fusion called deep Q-learning) to approximate Q-values in an unbounded environment (Wang et al., 2017).

Commonplace for deep neural networking is the multilayer perceptron (MLP) classifier, which articulates analysis as a classification problem, categorizing datapoints based on common attributes. MLP classifiers, comprised of several layers of nodes, classify data through a supervised learning process called backpropagation, wherein weights are repeatedly adjusted using a solving algorithm until loss is minimized (Lahiri et al., 2009). A sample MLP classifier with two hidden layers and three output classes can be seen below in *Figure 2*.

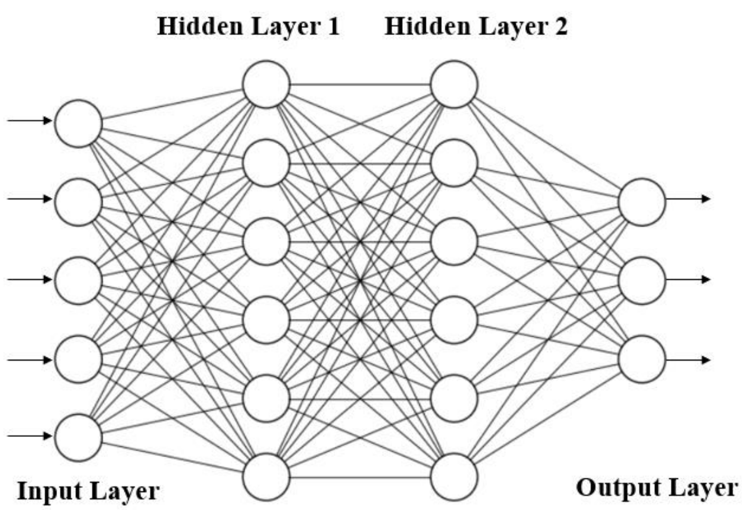

Figure 2. A Sample Multilayer Perceptron Classifier (Created by student researcher, 2021).

To aid machine learning-based programming, a multitude of software is available. As noted by Milmann & Avaizis (2011), Python is a free, high-level programming language with a vast selection of libraries. One of these libraries, Scikit-learn enables programmers to use Python's environment for the construction of machine learning algorithms, while also providing the building blocks for data analysis programs (Pedregosa et al., 2011).

Often used alongside ML systems are genetic algorithms (GA), which are evolution-based metaheuristics typically used for optimization problems (Katoch et al., 2020). In order to simulate pseudo-natural selection, genetic algorithms employ a repeating five-phase process, involving initialization, evaluation, selection, crossover, and mutation; this process can be seen below in *Figure 3*.

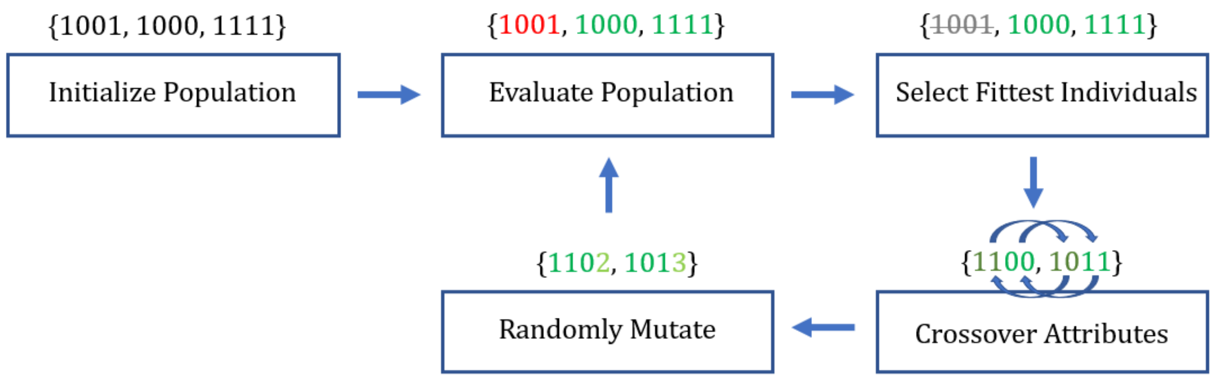

Figure 3. The Architecture of Genetic Algorithms (Created by student researcher, 2021).

In the most recent literature, genetic algorithms have been used to find the proper balance between elements required to effectively solve a variety of problems, ranging across politics (Kumar et al., 2014), engineering (Liang et al., 2020), medicine (Kim et al., 2021), and digital financial analytics (Wu et al., 2021). In this research, GA will be used to optimize technical indicators and preprocess them for integration into a deep Q-learning system. Until this point, the surrounding literature has largely void of research incorporating technical indicators into machine-learning-based systems, and completely void of research hybridizing technical indicators with reinforcement learning variants, such as deep Q-learning. The following experiment describes an attempt to fill this gap in research by integrating technical indicators into a deep Q-learning system through genetic algorithm optimization.

Statements of Purpose

1. Fuse deep Q-learning with technical indicators for the purpose of optimizing deep Q-learning-based trading.
2. Determine the relative value of technical indicators for the purpose of aiding analysts with the interpretation of the stock market.

Hypothesis

H_1: It was hypothesized that a deep Q-learning system, fundamentally fused with preprocessed technical indicators, would generate significantly higher average relative returns than a more traditional deep Q-learning approach and the benchmark buy-and-hold method.

Methods

The experiment discussed in this methodology was designed and conducted entirely by the student researcher. The stock market data used for the experiment ranges from January 1, 2015 to December 31, 2019. All data used was in the public domain, was sourced from Yahoo! Finance CSV database, and consists entirely of AMZN stock information. For the data processing, Python was selected as the coding language, and Spyder was selected as the integrated development environment. The following described methodology consists of four parts: data preprocessing using technical indicators, optimization of technical indicator weights using genetic algorithms, MLP classifier training using the optimized technical indicators, and simulating trading using the trained MLP classifier.

Data Preprocessing

Several indicators were integrated into a deep Q-system for use in this experiment. This integration required first preprocessing and then interpreting the market data using five popular technical indicators described below.

Exponential Moving Average (EMA)

The exponential moving average (EMA) is a weighted series of averages that applies higher weight to more current data. According to Fidelity (2019), if a stock is trending below an EMA, it can be interpreted as bearish (trending downward), whereas if a stock is trending above an EMA, it can be interpreted as bullish (trending upward). The EMA is calculated using the following formula. Note that a simple moving average is used for the first period in place of an EMA:

$$EMA_n = \frac{2(Current\ Price - EMA_{n-1})}{Period\ Length + 1} + EMA_{n-1} \tag{1}$$

Moving Average Convergence Divergence (MACD)

The moving average convergence divergence indicator (MACD) is a momentum indicator typically used to diagnose changes in trend (Fidelity, 2019). The MACD indicator consists of three components: the MACD line, the signal line, and a histogram. An approximate MACD can be calculated by subtracting the 12-period EMA by the 26-period EMA, the signal line can be calculated with a 9-period EMA of the MACD line, and the histogram can be calculated by subtracting the signal line from the MACD line. During trading, if the MACD line crosses zero in the upward direction, the stock is considered bullish, whereas if it crosses zero in the downward direction, the stock is considered bearish. Similarly, when the MACD crosses the signal line in the upward direction, the stock is considered bullish and when crossing the signal line downward, the stock is considered bearish. A sample MACD indicator can be seen below:

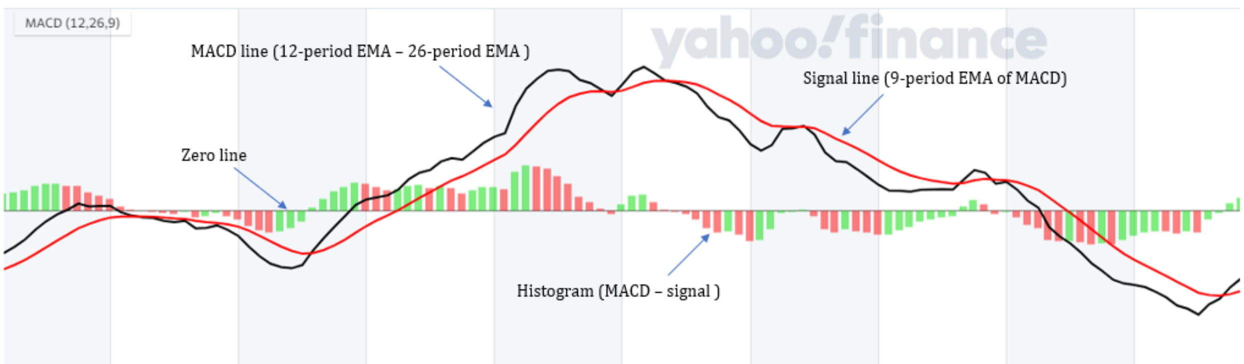

Figure 4. Diagram of MACD indicator on AMZN (Adapted from Yahoo Finance, 2021).

Money Flow Index (MFI)

The Money Flow Index (MFI) is a momentum indicator that determines the strength of a stock trend through the analysis of stock volume. The MFI oscillates with a range between 0 and 100, with a reading over 80 indicating overbought and a reading under 20 indicating oversold (Fidelity, 2019). The MFI is calculated using the below equation:

$$MFI = 100 - \frac{100}{1 + \left(\frac{14\ Period\ Positive\ Money\ Flow}{14\ Period\ Negative\ Money\ Flow}\right)} \tag{2}$$

Regarding the equation, the 14-Period Positive Money Flow equals the summation of upward trending days, wherein the typical price is higher than that of the previous day, while the

14-Period Negative Money Flow is determined by the sum of days wherein the typical price is lower than that of the previous day. Typical price is calculated using the formula depicted below in *Equation 3*.

$$Typical\ Price\ =\ \frac{Volume(High\ +\ Low\ +\ Close)}{3} \tag{3}$$

Relative Strength Index (RSI)

The Relative Strength Index (RSI) is a momentum oscillator that ranges between 0-100 and is used to identify reversals in trend. Typically, RSI indicates an overbought stock when reading above 70 and an oversold stock when reading below 30 (Fidelity, 2019). RSI is calculated using the following equation:

$$RSI\ =\ 100 - \frac{100}{(1 + \frac{Average\ of\ Positive\ Price\ Change}{Average\ of\ Negative\ Price\ Change})} \tag{4}$$

Bollinger Bands

Bollinger Bands are a three-line price envelope indicator consisting of a 20-period simple moving average (SMA), with an upper and a lower band, each located two standard deviations from the center SMA. The Bollinger Bands are indicative of volatility with the outer bands converging and diverging during periods of respective low and high volatility. The bands also often serve as barriers for stocks, with stocks tending to rebound off the bands.

Technical Indicator Interpretation

Once the raw data was transformed into indicator values, the market data were then interpreted according to the methods recommended by Fidelity, as seen below in *Table 1*. Note that Fidelity describes the interpretation of the MACD histogram by the qualitative appearance of its curve, which was reinterpreted quantitatively by the student researcher for ease of use in this experiment (see *Table 1*).

Table 1

Qualitative Technical Indicator Interpretation

Indicator	Reading	Interpretation
EMA (period=180)	Closing price > EMA	Bearish trend
	Closing price < EMA	Bullish trend
MACD	$\frac{dy}{dx}$(histogram) < 0 and $\frac{d^2y}{dx^2}$(histogram) < 0	Bearish trend
	$\frac{dy}{dx}$(histogram) > 0 and $\frac{d^2y}{dx^2}$(histogram) > 0	Bullish trend
	$\frac{d^2y}{dx^2}$(histogram) < 0 and 0.5 > $\frac{dy}{dx}$(histogram) > -0.5	Bearish trend
	$\frac{d^2y}{dx^2}$(histogram) > 0 and 0.5 > $\frac{dy}{dx}$(histogram) > -0.5	Bullish trend
MFI	MFI > 80	Bearish trend
	MFI < 20	Bullish trend
Bollinger Bands	Closing price > upper band	Bullish trend
	Closing price < lower band	Bearish trend
RSI	RSI > 60	Bearish trend
	RSI < 40	Bullish trend

Following interpretation, the trend reading was normalized, with “0” representing a bearish trend, “1” representing bullish trend, and “0.5” representing a neutral trend. These normalized data values were then weighted to return trend prediction labels, with the weights being optimized using a genetic algorithm (*Figure 5*), whose fitness function is depicted below in *Equation 5*. Regarding the algorithm, a relatively low number of generations per population (50) was selected. This choice as well as the decision to implement a normalization function were intended to compensate for the algorithm’s tendency to converge on local minima.

$$Fitness = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{dy}{dx}(price_{n+1}) - W_1T_{EMA_{180}} - W_1T_{MACD} - W_1T_{MFI} - W_1T_{Bollinger} - W_1T_{RSI}\right| \tag{5}$$

Where *T* represents predicted trend at period *n* and *W* represents variable weight.

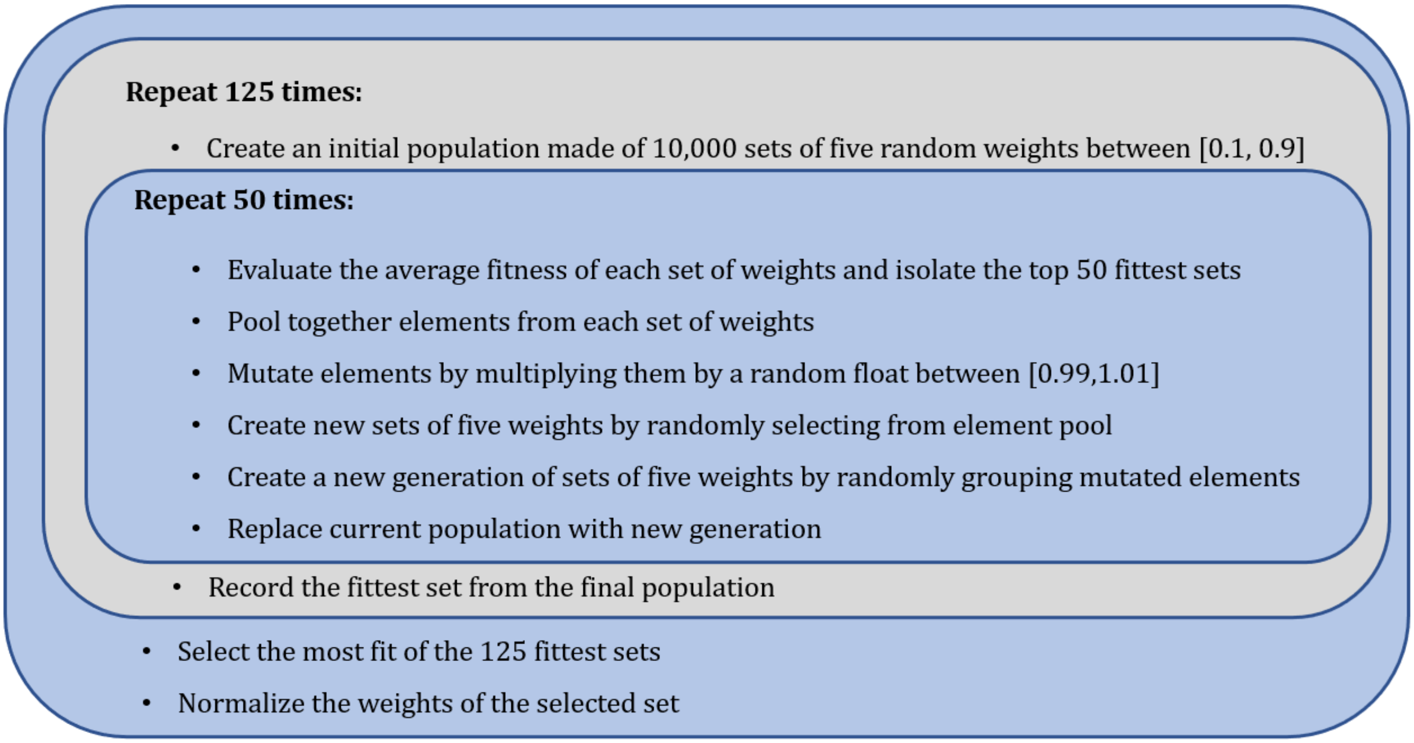

Figure 5. Visual representation of the genetic algorithm (Created by student researcher, 2021).

The normalization formula referred to in *Figure 5* was created by the student researcher and is a modified version of a sigmoid function that transformed the weights into floats between 0.1 and 0.9, while maintaining their distinctiveness. This formula can be seen below in *Equation 6*.

$$N(x) = \frac{1}{1 + \frac{0.05}{x}} \tag{6}$$

Multi-layer Perceptron Classifier Training

Following the derivation of the trend prediction labels with the genetic algorithm, the internal reward function was determined, allowing for the rest of the deep Q-learning process to be completed. Both the training and simulation phases were conducted on AMZN stock and used the Scikit Learn framework for the neural networking. During both training and simulated trading, the state vector was comprised of 150 past prices, the current budget, and the current number of shares owned, while the traditional investment decisions of buy, hold, and sell were selected as potential actions. As for the rewards functions, two distinct formulas were used to weigh reward and determine Q. For the benchmark trials, *Equation 7* was used, whereas *Equation 8* was used for the hybrid TA approach. Once Q was determined, A Q-matrix was constructed using *Equation 9*.

$$Q_1 = 0.5 * \left(\Delta price_{period=10} + \Delta price_{period=1}\right) \tag{7}$$

$$Q_2 = W_1 T_{EMA_{180}} + W_1 T_{MACD} + W_1 T_{MFI} + W_1 T_{Bollinger} + W_1 T_{RSI} \tag{8}$$

$$Q_{(s,a_1,a_2,a_3)} = \left|\frac{1}{Q-0}\right|, \left|\frac{1}{Q-0.5}\right|, \left|\frac{1}{Q-1}\right| \tag{9}$$

With $Q_{(s,\max(a))}$ representing the optimal action *a* at state *s*, an optimal decision policy was derived over the course of 100 simulated investment trials. Roughly ten percent of the actions made in training were random, as to ensure the complete exploration of the sample space. Once an optimal decision policy was derived, the corresponding optimal state-action pairs were inputted as features and labels into the neural network.

Multi-layer Perceptron Classifier Testing

The final simulation that served to test the hybrid deep Q-learning strategy operated on the last 20 percent of the data and traded with an initial budget of $10,000. For each closing price, a state vector composed of the recent price history, the number of shares owned, and the current budget was inputted into the neural network. Each inputted state returned a corresponding optimal action which updated the environment and produced the next state. After each iteration through the designated testing dataset, the final portfolio value was recorded for interpretation.

Results

The simulation was conducted a total of 101 times: with 50 trials devoted to the standard deep Q-learning method, 50 trials devoted to the hybrid method, and a single trial devoted to the benchmark buy-and-hold method. Only a single trial was necessary for the buy-and-hold method as no variance was possible for this strategy. For each method tested, average training returns, average testing returns, training standard deviation, and testing standard deviation were reported. The results of the simulations can be seen below in *Table 2*. Additionally, the distributions of simulated returns during testing can be seen below in *Figure 6* and *Figure 7*, with both distributions featuring a positive skew.

Table 2

Comparison of Trading Performance

	Average Simulated Returns (Training)	Average Simulated Returns (Testing)	Simulated Returns Std Dev (Training)	Simulated Returns Std Dev (Testing)
Buy-and-hold Method	N/A	$9544.00	N/A	$0.00
Standard deep Q-learning Method	$10815.26	$10469.57	$2816.84	$856.61
Hybrid deep Q-learning Method	$11065.57	$10933.91	$1785.29	$397.76

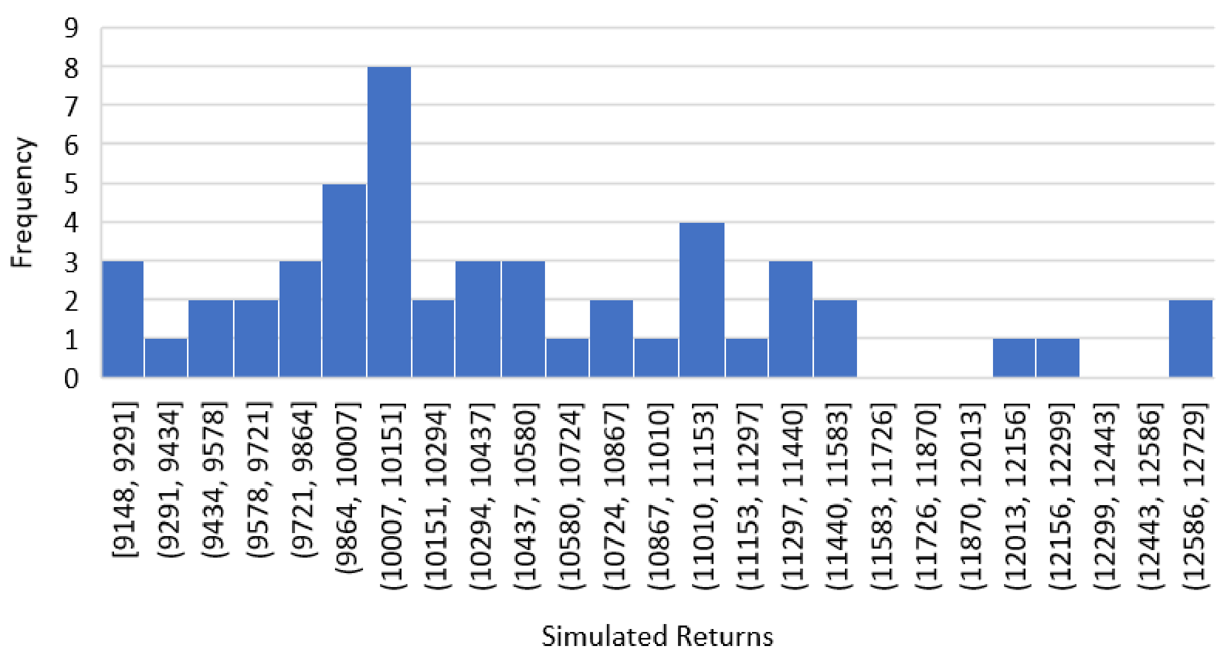

Figure 6. The distribution of simulated returns (standard deep Q-learning method). (Created by student researcher, 2021).

Figure 7. The distribution of simulated returns (hybrid method). (Created by student researcher, 2021).

The performance of each strategy was evaluated based on average simulated returns and return variance. The hybrid strategy proposed in this experiment produced consistently higher returns than the standard deep Q-learning method and the buy-and-hold method. Additionally, the hybrid strategy produced consistently lower variance compared to the standard deep Q-learning method and the buy-and-hold method.

Regarding the final state of the genetic algorithm, the optimal relative weights for the tested technical indicators can be seen below in *Table 3*.

Table 3

Derived Optimal Relative Weights

Technical Indicator	RSI	MACD	$EMA_{period=180}$	MFI	Bollinger Bands
Derived Optimal Weight	0.633	0.572	0.574	0.962	0.857

Discussion

This experiment intended to optimize a deep Q-learning system for stock market trading through the integration of technical indicators. The GA process involved in the integration was also expected to reveal the optimal weights of the tested market indicators. The proposed hypothesis of this research was that a deep Q-learning system, fundamentally fused with preprocessed technical indicators, would generate significantly higher average relative returns than a more traditional deep Q-learning approach and the benchmark buy-and-hold method. This hypothesis was supported by the results as the average simulated returns in testing for the hybrid deep Q-learning method were statistically significantly higher ($p=0.000851839$) than those returns from the buy-and-hold approach, as well as those from the standard Q-learning approach. From these results, it can be derived that preprocessed market indicators, when they are properly weighed and substituted for the traditional long-term and short-term reward functions, can fundamentally increase the profitability of deep Q-learning-based trading. From the higher relative profits produced by the hybrid model, it is also reasonable to assume that the the GA process was effective in optimizing technical indicator weights, as the hybrid method was trained more successfully with technical indicators than was the traditional method using the Bellman Equation.

Considering that the GA process produced weights effective in training the hybrid model, these resultant relative weights could be used in future experiments by researchers that wish to weigh or prioritize the indicators that they are applying to stock market data. Additionally, these relative weights could be applied outside of automated endeavors by any investors who wish to gauge which indicators are most indicative of future stock direction. The relative value of each indicator tested corresponds to its weight in *Table 3*, with the indicators of the highest weight bearing the most relative value. Considering this, the MFI has been found by this experiment to be the indicator most suggestive of future stock direction, as it has the highest relative weight (0.962). This high value can likely be attributed to the fact that the MFI analyzes stock volume in addition to price, which in turn suggests that volume, as a contributing measurement to MFI, is also indicative of future stock price, so long as it is properly interpreted.

As for the simulated returns graphs seen in *Figure 6* and *Figure 7*, both distributions feature a positive skew. Positive skewness in the returns on investments is commonly interpreted in financial theory as being indicative of an investment strategy involving frequent minor losses paired with infrequent substantial gains. Considering that the automated trading of this experiment produced an investment strategy mirrored by real investors, automated deep Q-learning-based trading, as conducted during this experiment, may have real-world financial applications, especially as an additional tool in trend prediction.

It should be noted that the simulation was conducted only using data gathered from AMZN stock. However, despite the overall positive trend of AMZN, the experimental results are not expected to be inflated, as the testing of both the hybrid and standard deep Q-learning models was conducted over an overall negatively trending period for AMZN stock.

Conclusion

For this research, a deep Q-learning trading system was successfully constructed and optimized by fusing it with technical indicators. This optimal system proved to be more consistent and statistically more profitable than its more traditional counterpart. The hypothesis of this experiment was supported by the results as the hybrid deep Q-learning system, fundamentally fused with preprocessed technical indicators, generated significantly higher average relative returns than both the more traditional deep Q-learning approach and the benchmark buy-and-hold method.

An ancillary product of the genetic algorithm used in this research was a resultant relative value for the tested technical indicators. From these weights, it was deduced that the MFI indicator was valued most highly by the hybrid trading system. As such, the MFI indicator may have significant value to real-world analysts keen on deciphering the market.

A novel finding of this experiment was observable in the skew of the average simulated returns. Considering that the returns were skewed right, it is reasonable to assume that the system favored the strategy of frequent investments, many of them losses, with a select number of them being highly profitable. As this strategy far outperformed the benchmark buy-and-hold method, it also likely has potential to be used by real investors, and these results should invite the embracement of similar strategies.

Further work should be done integrating technical indicators with machine learning. This experiment explored exclusively deep Q-learning, yet other variants of machine learning may produce more profitable investment strategies given the integration of properly preprocessed technical indicators. Additionally, further work should be done applying technical indicators outside the realm of stock market price prediction, as technical analysis likely has the potential to aid the automated interpretation of many convoluted datasets across many different fields.

References

Amazon.com, Inc. (AMZN) Interactive Stock Chart - Yahoo Finance. (n.d.). Finance.yahoo.com. https://finance.yahoo.com/quote/AMZN/chart?p=AMZN

Cumming, J., Alrajeh, D., & Dickens, L. (2015). An investigation into the use of reinforcement learning techniques within the algorithmic trading domain. Imperial College London: London, UK.

Dash, R., & Dash, P. K. (2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. The Journal of Finance and Data Science, 2(1), 42-57.

Elbialy, B. A. (2019). The Effect of Using Technical and Fundamental Analysis on the Effectiveness of Investment Decisions of Traders on the Egyptian Stock Exchange. International Journal of Applied Engineering Research, 14(24), 4492-4501.

García Cuenca, L., Puertas, E., Fernandez Andrés, J., & Aliane, N. (2019). Autonomous driving in roundabout maneuvers using reinforcement learning with Q-learning. Electronics, 8(12), 1536.

Jakpar, S., Tinggi, M., Tak, A. H., & Chong, W. Y. (2018). Fundamental Analysis VS Technical analysis: The Comparison of Two Analysis in Malaysia Stock Market. UNIMAS Review of Accounting and Finance, 2(1).

Katoch, S., Chauhan, S. S., & Kumar, V. (2021). A review on genetic algorithm: past, present, and future. Multimedia Tools and Applications, 80(5), 8091-8126.

Kim, C., Batra, R., Chen, L., Tran, H., & Ramprasad, R. (2021). Polymer design using genetic algorithm and machine learning. Computational Materials Science, 186, 110067.

Kumar V, Chhabra JK, Kumar D (2014) Parameter adaptive harmony search algorithm for unimodal and multimodal optimization problems. J Comput Sci 5(2):144–155

Lagoudakis, M. G., & Parr, R. (2003). Reinforcement learning as classification: Leveraging modern classifiers. In Proceedings of the 20th International Conference on Machine Learning (ICML-03) (pp. 424-431).

Lahiri, S. K., & Ghanta, K. C. (2009). Artificial neural network model with the parameter tuning assisted by a differential evolution technique: The study of the hold up of the slurry flow in a pipeline. Chemical Industry and Chemical Engineering Quarterly/CICEQ, 15(2), 103-117.

Li, D., Wang, Y., Madden, A., Ding, Y., Tang, J., Sun, G. G., ... & Zhou, E. (2019). Analyzing stock market trends using social media user moods and social influence. *Journal of the Association for Information Science and Technology*, *70*(9), 1000-1013.

Liang, H., Zou, J., Zuo, K., & Khan, M. J. (2020). An improved genetic algorithm optimization fuzzy controller applied to the wellhead back pressure control system. Mechanical Systems and Signal Processing, 142, 106708.

MACD Indicator | What it's saying about stocks now | Fidelity. (n.d.). Www.fidelity.com. https://www.fidelity.com/viewpoints/active-investor/how-to-use-macd

Milmann, K. J., & Avaizis, M. (2011). Scientific python. Computing in Science & Engineering, 11.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602.

Mosavi, A., Ghamisi, P., Faghan, Y., & Duan, P. (2020). Comprehensive Review of Deep Reinforcement Learning Methods and Applications in Economics. arXiv preprint arXiv:2004.01509.

Nazário, R. T. F., e Silva, J. L., Sobreiro, V. A., & Kimura, H. (2017). A literature review of technical analysis on stock markets. *The Quarterly Review of Economics and Finance*, *66*, 115-126.

Pahwa, N. S., Khalfay, N., Soni, V., & Vora, D. (2017). Stock prediction using machine learning a review paper. International Journal of Computer Applications, 163(5), 36-43.

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert systems with applications, 42(1), 259-268.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.

Sah, S. (2019). Machine Learning: A Review of Learning Types.

Saldivar, F., & Ortiz, M. Stock Market Price Prediction Using Various Machine Learning Approaches. 7

Wang, J. L., & Chan, S. H. (2007). Stock market trading rule discovery using pattern recognition and technical analysis. Expert Systems with Applications, 33(2), 304-315.

Wang, Y., Wang, D., Zhang, S., Feng, Y., Li, S., & Zhou, Q. (2017). Deep Q-trading. cslt. riit. tsinghua. edu. cn.

Watkins, C. J., & Dayan, P. (1992). Q-learning. *Machine learning*, *8*(3-4), 279-292.

What Are Bollinger Bands? - Fidelity. (2019). Fidelity.com. https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/bollinger-bands

What Is EMA?- Exponential Moving Average - Fidelity. (n.d.). Www.fidelity.com. https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/ema

What Is MACD? - Moving Average Convergence/Divergence - Fidelity. (n.d.). Www.fidelity.com. https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/macd

What is RSI? - Relative Strength Index - Fidelity. (2019). Fidelity.com. https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/RSI

What Is The Money Flow Index? - Fidelity. (n.d.). Www.fidelity.com. https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/MFI

Wu, Y., Li, X., Liu, Q., & Tong, G. (2021). The Analysis of Credit Risks in Agricultural Supply Chain Finance Assessment Model Based on Genetic Algorithm and Backpropagation Neural Network. Computational Economics, 1-24.

Xiong, Z., Liu, X. Y., Zhong, S., Yang, H., & Walid, A. (2018). Practical deep reinforcement learning approach for stock trading. arXiv preprint arXiv:1811.07522.

Zhang, C. (2019). Sentiment Analysis and Deep Reinforcement Learning for Algorithmic Trading.

National Junior Science & Humanities Symposium
Statement of Outside Assistance

(Student finalists presenting their research paper at the Regional and National symposium must complete this form and submit with the final research Paper.)

Name: Matthew Kahn

Regional Symposium: Upstate New York

Title of Paper: The Integration of Technical Indicators into a Deep Q-Learning System to Increase

1) Please explain your role in the development of the project idea.

I developed the project idea completely independently, incorperating concepts from old research projects and independent financial investigation.

2) What steps led you to formulate your research question? --or—What steps led you to develop the design for your engineering project?

Last year, I conducted research involving deep Q-learning, and outside of Science Research, I had taken an interest in stock market technical analytics. The idea for my project came when I tried to fuse deep Q-learning with financial analytics tools and found the literature to be largely absent of similar research. The design was then relatively simple as I already had experience coding stock market

3) Where did you conduct the major part of your work? (i.e. home, school, or other institutional setting – university lab, medical center, etc.)

I conducted my work entirely at home.

4) Describe the assistance that you received throughout the project.

My mentor introduced me to the statsitical tests necessary for the experimental analysis. Outisde of this, the project was entirely my own doing.

5) If you worked in an institutional setting, describe your role on the team.

N/A

6) What role did each person play in the research investigation?

N/A (only one student researcher)

7) Describe what parts of the research you did on your own and what parts where you received help. (i.e. literature search, hypothesis, experimental design, use of special equipment, gathering data, evaluation of data, statistical analysis, conclusions, and preparation of written report (abstract and/or paper).

My mentor introduced me to certain statistical tests such as the T-test, and reccomended statistical tests for me to use for my statistical analysis. Aside from this, all other parts of the research, including the written report, were completed entirely on my own.

8) If this research is a continuation of an investigation that was previously submitted to a regional JSHS, describe how you have expanded your investigation.

N/A

Comments by teacher and/or supervising mentor on the students' individual contributions to the research investigation or engineering/computer science project:

Matthew was the sole contributer to his research. He was responsible for all aspects of the design and implementation of his research plan.

Statement by the teacher or supervising mentor acknowledging that the student conducted the research in accordance with proper procedures and protocols for the conduct of animal research or human research. Projects which were conducted without proper supervision will be disqualified from both regional and National competition. Further guidelines may be found at http://www.jshs.org

- Research activities involving non-human vertebrates or human subjects must be submitted for IRB review prior to the conduct of the research.
- Research activities involving vertebrate animals must be conducted in compliance with local, state, and federal guidelines for the humane and ethical treatment of animals in the conduct of the research.

Please have the supervising teacher and/or supervising scientist sign below. If you did the work without a teacher or supervising scientist, you will need a signature from your parent and a brief description of their role in the research.

11/24/21 Matthew Kahn Somers High School

Date | Signature of Student (Required) | High School

11/24/21 William E Maelia Somers High School

Date | Signature of Teacher | High School

Date | Name of Supervising Scientist | Title of Supervising Scientist

Institution of Supervising Scientist

Signature of Supervising Scientist