International Journal of Innovative Engineering Applications

Journal homepage:https://dergipark.org.tr/ijiea

GENETIC ALGORITHM BASED APPROACH FOR ALGORITHMIC TRADING IN FINANCIAL MARKETS

***Yunus Santur**[*1]*, ***Mustafa Ulas**[2]*, ***Murat Karabatak**[1]*

[1]*Firat University, Faculty of Technology, Department of Software Engineering, Elazig, Turkey*
[2]*Firat University, Faculty of Engineering, Department of Software Engineering, Elazig, Turkey*

Abstract

Original scientific paper

Software that enables realtime buy and sell transactions in financial markets according to predetermined conditions is called algorithmic trading. When developing algorithmic trading robots, indicators used in technical analysis are generally used. For the strategy selection of the robot, a process called Backtest is performed on the historical time series. The purpose of the Backtest process is the process of obtaining and interpreting values such as the number of successful/unsuccessful trades, the portfolio cash value after the commission to be paid to the intermediary institution, the profit factor and the sharpe ratio. The biggest disadvantage in this process is the selection of the appropriate stock, period, indicator and their parameters. Linear programming approaches are mostly used in the selection of these parameters that optimize the Backtest process optimally. However, according to the strategy to be used, the coding of these algorithms can have a linear, quadratic or polynomial complexity. This requires more long testing times for investors and algorithmic robot developers. Genetic algorithm-based approaches inspired by nature, on the other hand, converge to the optimal solution with much less iteration and require less processing power and time. In this study, a genetic programming-based approach is proposed for the selection of optimal conditions in algorithmic trading. In the experimental studies section, it has been seen that the use of traditional and genetic algorithm-based approaches in algorithmic trading operations has advantages when comparing complexity.

Keywords: *Algorithmic trading, genetic algorithm, optimization.*

FİNANSAL PİYASALARDA ALGORİTMİK TİCARET İÇİN GENETİK ALGORİTMA TEMELİ YAKLAŞIM

Özet

Orijinal bilimsel makale

Finansal piyasalarda önceden belirlenmiş koşullara göre anlık al sat işlemlerinin yapılmasını sağlayan yazılımlara algoritmik ticaret denilmektedir. Algoritmik işlem robotları geliştirilirken genellikle teknik analizde kullanılan göstergeler kullanılmaktadır. Robotun strateji seçimi için geçmiş veriler üzerinde Backtest adı verilen işlem gerçekleştirilmektedir. Backtest işleminin amacı gerçekleştirilen başarılı/başarısız ticaret sayısı, aracı kuruma ödenecek komisyon sonrası portföy kasa değeri, kar faktörü ve sharpe oranı gibi değerlerin elde edilerek yorumlanması işlemidir. Bu süreçte en büyük dezavantaj uygun stok, periyot, indikatör ve bunlara ait parametrelerin seçimidir. Backtest işlemini optimal olarak en iyileyen bu parametrelerin seçiminde çoğunlukla doğrusal programlama yaklaşımları kullanılmaktadır. Ancak kullanılacak stratejiye göre bu algoritmaların kodlanması lineer bir karmaşıklıktan, quadratic veya polynomial karmaşıklığa sahip olabilmektedir. Bu durum yatırımcılar ve algoritmik robot geliştiriciler için uzun test süreleri gerektirmektedir. Doğadan esinlenerek geliştirilen genetik algoritma tabanlı yaklaşımlar ise çok daha az iterasyon ile optimal çözüme yakınsayarak, daha az işlem gücü ve zaman gerektirmektedir. Bu çalışmada algoritmik ticarette optimal koşulların seçimi için genetik programlama tabanlı bir yaklaşım önerilmiştir. Deneysel çalışmalar bölümünde, geleneksel ve genetik algoritma tabanlı yaklaşımların karmaşıklık, benchmark ve Backtest sonuçları karşılaştırıldığında algoritmik ticaret işlemlerinde kullanılmasının avantajlara sahip olduğu görülmüştür.

Keywords: *Algoritmik ticaret, genetik programlama, optimizasyon.*

1 Introduction

Transactions in stocks, crypto-assets and futures markets are handled in two classes as manual and robotic. Manual transactions include dividend investing, long-term investment, short and long-term investments in line with company/sector expectations. It is known that mostly manual operations are performed around the world. However, it is known that the number of investors making algorithmic transactions has increased in recent years. A

*Corresponding author.

E-mail address: ysantur@firat.edu.tr (Y. Santur)

Received 18 September 2022; Received in revised form 14 May 2023; Accepted 28 August 2023

 Doi: https://doi.org/10.46460/ijiea.1176888

significant size of the transaction volume in financial markets is performed by algorithmic robots. Unlike manual operations, algorithmic operations are performed automatically by software that reads news/text and/or uses technical analysis indicators on temporal series [1-5].

Figure 1. Financial time series data a) OHLC notation in financial time series b) Financial time series and volume.

$$candle = \begin{cases} Bullish, if\ Close > Open \\ Bearish, otherwise \end{cases} \quad (1)$$

It is known that there are hundreds of indicators that have entered the literature in technical analysis and have been developed as open or close sourced.

These indicators are obtained by using the open, close, low, high and transaction volume information of the financial asset within the relevant period. In the Figure 1, the time series of the financial asset is shown by using the Volume amount as a bar chart in the lower part of the price chart, and the candle charts consolidating the O:Open, C:Close, L:Low and H:High price values in a single visual at the top. While creating candlestick charts, a single image can be obtained from the OHLC data, which are the four values mentioned [6].

Eq.(1) is used when naming candlestick charts. Although they express Bullish or Bearish in the simplest terms, pattern recognition and the development of pattern-based approaches in the literature is a separate study. Because it is known that there are more than 100 candle patterns. Dozens of different patterns are formed by the ratio of body and shadows, and the combination of these patterns in pairs and triples [7].

In Figure 2, the most known and widely used indicators obtained from time series are given. The formulas for these and the Buy/Sell strategies that can be created are given in Table 1. The indicators in Table 1 are expressed as : Simple moving average (SMA), Exponential Moving Average (EMA), Moving Average Convergence/Divergence (MACD), Relative Strength Index (RSI) and Momentum (Mom), respectively [8] .

The SMA given in Eq.(2) represents the moving average of the selected n periods. It is assumed, above the selected period trend is bullish, below the selected period trend is bearish. In the equation, C represents the closing values, but any of the OHLC can be chosen according to the strategy.

Table 1. Commonly used indicators and trading strategies.

Indicator	formula	Strategy	Eq.
SMA	$SMA = {}^{\sum C_i}/_{n}$	• The price cuts the average up • The shorter of the averages of two different periods cuts the longer one up.	(2)
EMA	$EMA_t = \begin{cases} C_0, if\ t = 0 \\ \alpha + (1-\alpha)EMA_{t-1}, if\ t > 0 \end{cases}$		(3)
MACD	$MACD_{line} = EMA(12) - EMA(26)$ $Signal_{line} = EMA(MACD_{line}, 9)$	• MACD value cuts up Signal value	(4)
RSI	$RSI = 100 - \frac{100}{1 + {}^{\mu_{Gain}}/_{\mu_{Loss}}}$	• Intersection of RSI value with 30/70 threshold values • RSI value cutting its average up	(5)
mom	$MOM = 100 * \frac{C_i}{C_{i-n}}$	• The intersection of the MOM value with the threshold value	(6)

The EMA given in Eq. (3), unlike the SMA, does not weight the bars equally, it is calculated recursively according to the values of the previous bars. Thus, it is more sensitive to recent price movements. The MACD given in Eq. (4) was used for the Japanese markets in the early years. It takes the difference of the EMA value in two different periods, which is also used in the SMA and EMA strategy, while Signal provides the smoothing value in the default 9 periods. Since the Japanese markets are open six days a week, the strategy is formed by calculating the difference of the 2-week buying momentum (12) to the one-month buying momentum (26) and its intersection with its 9-day softened value. Although the default period is calculated based on daily values, different periods can be selected.

The RSI given in Eq.(5) is a common indicator used for excess, normalized to the 0-100 range. Below 30 means oversold and over 70 means overbought. As a general acceptance, a position is taken assuming that if it goes above 30 again, it is bullish, and if it goes below 70 when it is above 70, it will be a bearish trend.

Another strategy is to use the intersection of the RSI value with its average according to the n periods to be selected, as a common option. The Mom given in Eq. (6) represents the percentage change according to the selected n periods. The main motivation of the strategy used in the Mom indicator is to avoid horizontal market noise.

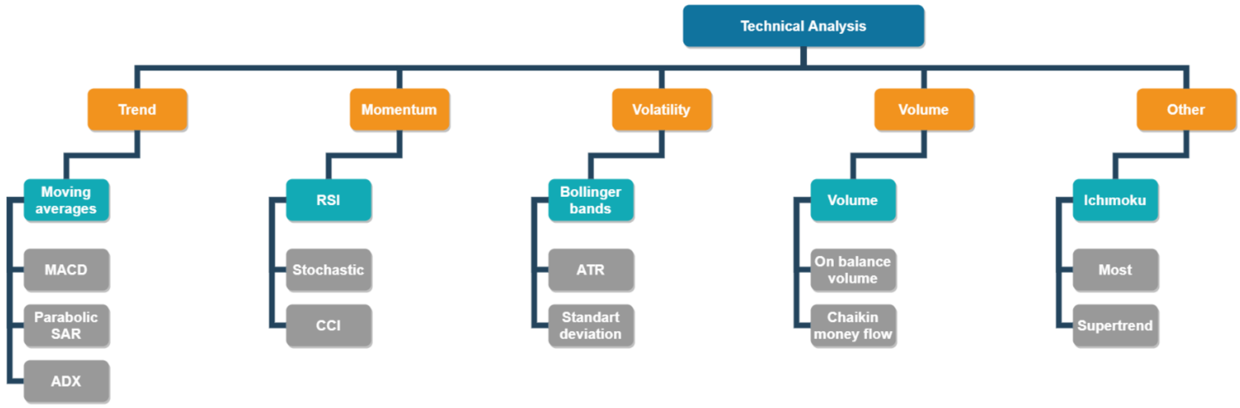


Figure 2. Technical analysis indicators.

2 Materials and Methods

In this study, a genetic algorithm-based approach is proposed for the parameter selection of indicators, periods and indicators in the algorithmic robot development process. The data terminals used in this area handle the parameter selection optimization process with linear programming by using time series data in the backtest process. For example, the MACD indicator takes 3 parameter data as input. Considering the selection of periods such as day, 4-hour, hourly, combining with other indicators and stock selection, optimizing with nested loops has quadratic or polynomial complexity. This situation causes a disadvantage in the selection of hundreds of stocks, the period to be selected, the indicator to be used and the periods of the indicators. Genetic algorithms inspired by nature offer a much more effective solution to the optimization problem. In the continuation of the article, information about the data set and the proposed approach, data and experimental studies are presented.

2.1 Dataset

In the study, data sets of financial assets given in Table.2 were obtained from Yahoo finance platform and used [9].

Table 2. Commonly used indicators and trading strategies.

Data set	Explanation	Date range
Bist100	Borsa Istanbul national 100 index	27.08.2020 22.08.2022

Yahoo finance platform provides OHLCV (OHLC + V:Volume) data for almost all financial assets in the desired period and date range free of charge in .csv format. The screenshot obtained from the raw data set in the application environment is shown in Figure 3.

Date	Open	High	Low	Close	Volume
2022-03-25	2171.5073	2187.1497	2165.6291	2175.5000	2.113966e+09
2022-04-15	2513.7488	2536.2945	2510.1943	2494.3999	3.444524e+09
2022-04-20	2525.8999	2536.5000	2502.3999	2525.8999	4.118388e+09
2020-12-04	1330.2004	1336.4939	1324.8061	1330.9000	5.295296e+09
2022-06-29	2314.6930	2316.2716	2230.5594	2402.0000	2.696174e+09

Figure 3. Dataset screenshot.

2.2 Backtest Optimization.

Stock, period and the code structure of the traditional approach used in data terminals for parameter selection of indicators is given in Figure 4. In this approach, nested iterations are performed for the selected range as much as the total number of parameters.

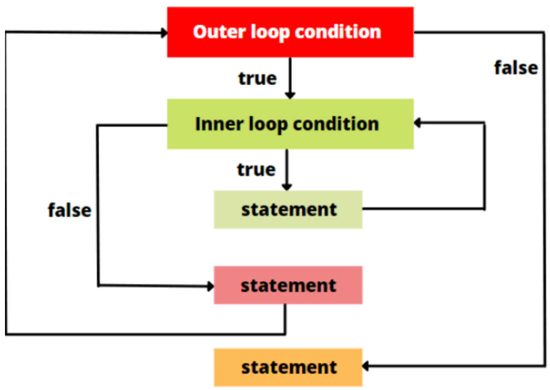


Figure 4. Optimizing with nested loops.

This approach increases the Big-O complexity exponentially according to the number of parameters to be selected [10]. Although this complexity increases, since almost all of the transactions performed are simple conventional (statistical) calculations, the processing computations required is mostly simple. However, the main problem here is the occurrence of overfitting as a result of optimization. Financial time series are chaotic and cannot be expected to fluctuate in the same way in the future. Therefore, it cannot be guaranteed that the values that provide the best optimization will yield good returns in the future. For this reason, the optimization phase is also supported by Walk Forward Analysis (WFA) or Monte Carlo simulation (MCS). All these are the main disadvantages of traditional Backtest optimization [11, 12].

2.3 Genetic Algorithm Based Backtest Parameter Selector

Genetic algorithms are approaches that are developed by observing nature and realize the optimization with an evolutionary mechanism [13-15]. Genetic algorithms are suitable for Backtest optimization by their nature. It is an effective method that can provide convergence to the most optimal parameters without the need for process complexity and later approaches such as WFA, MCS and without overfitting.

The proposed approach in this study is given in Figure 5. The approach, which takes the time series as input, makes the parameter selections that optimize the loss function of the network for the selected indicators adaptive. The proposed method can be considered recursively, as well as selecting the most appropriate parameter, combining the most appropriate input value (Open, Close, Low, High) values, the most appropriate period (daily, hourly, weekly) and the most compatible indicators.

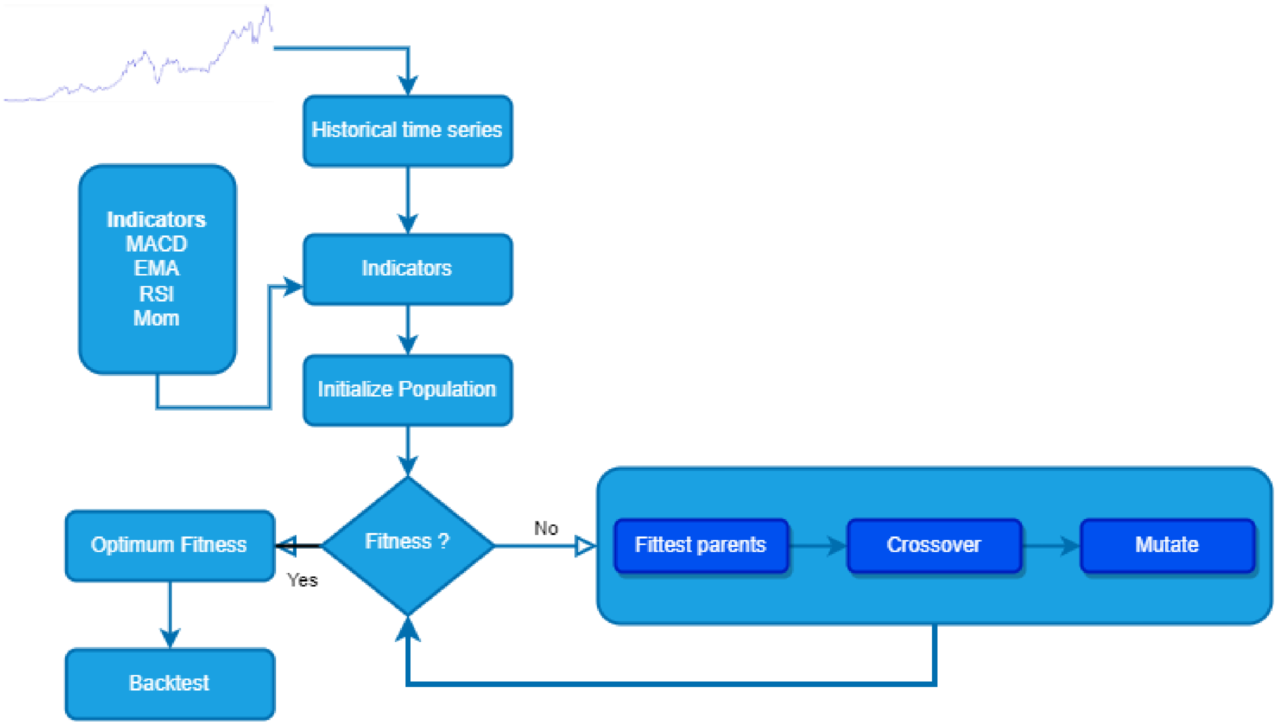


Figure 5. Proposed method.

3 Experiments and Results

Experimental studies consist of 500 pieces of data consisting of daily closing values of approximately 2 years in the Bist100 index. Backtest procedure was used to compare the results [16, 17]. Experimental studies were carried out using Python programming language and pandas, backtrader libraries [18-20]. In the backtest process, slippage and commission are taken into account as (1/10000). In all studies, the initial safe value was assumed to be 10,000. Buy/sell signals obtained at bar closings are shown in Figure 6.

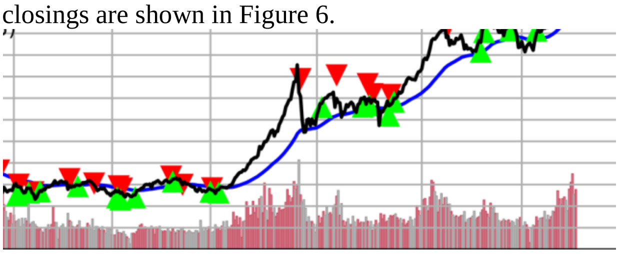


Figure 6. Buy/sell signals.

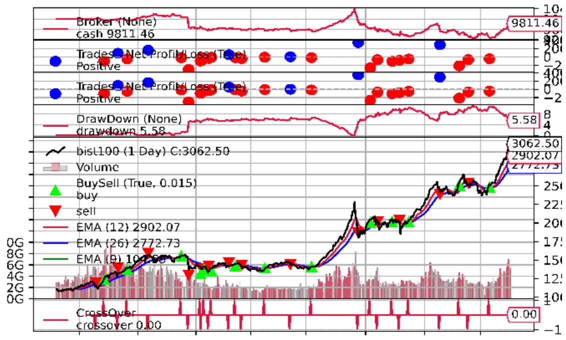
In the aggregated results given in Table 2, better results were obtained with genetic algorithms for all indicators. With the proposed approach, it was observed that the portfolio value increased by 28%. Another finding is that the MACD, EMA and RSI indicators work better. When the proposed approach and conventional-based approaches are compared, the portfolio returns increased between 2% and 30% thanks to the proposed approach.

All strategies are designed in two ways. In other words, when the Long signal occurs, if there is a Short position, it buys by taking profit or by closing it with a stop loss, and on the contrary, it closes the Long positions and takes a Short position. For this purpose, positions are opened twice as much as the previous one each time. In the executed Backtest strategy, it can be assumed that 1 lot is purchased at a time or that a cumulative purchase is made at the ratio of the portfolio. In this study, the second one, the cumulative portfolio, was preferred. Better results are shown as bold in Table 2.

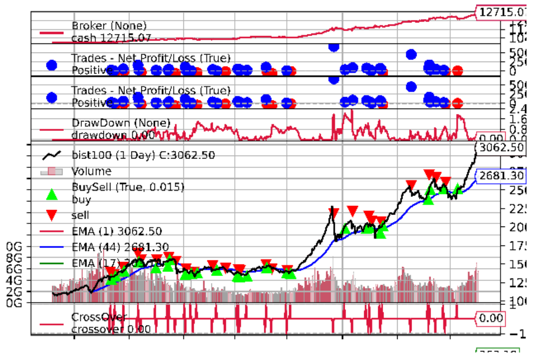
In Figure 7 below, the trades and portfolio gains obtained with the conventional and proposed approach for MACD, EMA and RSI in conventional Table 3 from the backtester library are shown collectively.

Table 3. Experimental studies.

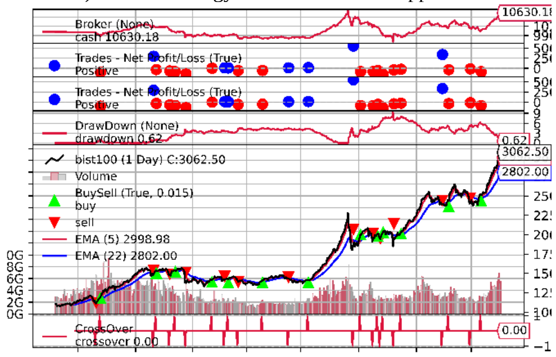
Indicator	Conventional Approach		Recommended approach		Strategy
	Default parameters	Portfolio return (%)	Obtained parameters	Portfolio return (%)	
MACD	12 – 26 – 9	-1.9%	3-41-2	**28%**	Long trade when MACD crosses above signal line up
C - EMA	1 - 8	16%	3-34	**18%**	Long trade when close cross above 8-ema
EMA	5 - 22	6%	3-34	**18%**	Long trade when the fastest EMA crosses above slowest EMA
RSI	5 - 14	13%	7 - 14	**19%**	Long trade when RSI recursively cross above its moving average up
Mom	1.0	15%	.94	**17%**	Long trade when Mom exceeds percentile



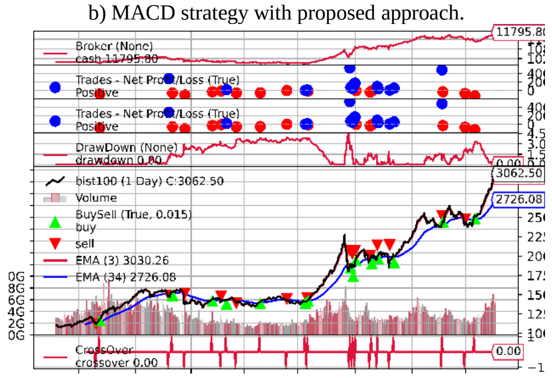
a) MACD strategy with conventional approach.



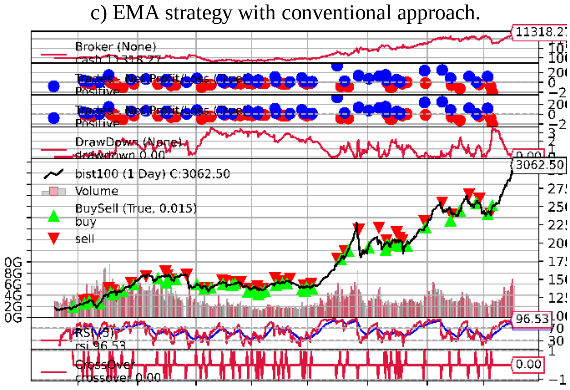
b) MACD strategy with proposed approach.



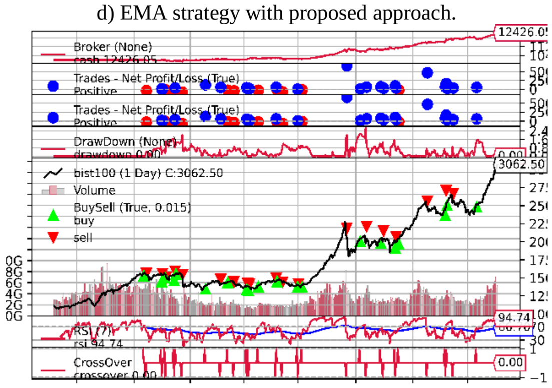
c) EMA strategy with conventional approach.



d) EMA strategy with proposed approach.



e) RSI strategy with conventional approach.



f) RSI strategy with proposed approach.

Figure 7. Backtest results

In Figure 8, the average and optimum suitability values typically obtained in experimental studies are shown graphically. When nested loops are used in the optimization process based on the backtest process, much more iterations work. Genetic algorithms, on the other hand, are very suitable for this problem by their nature, they can converge to the optimum solution by exploring the problem space with much less iteration.

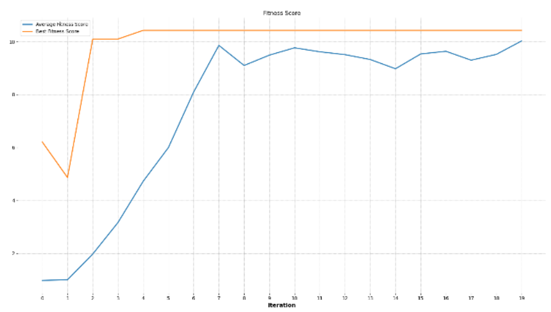


Figure 8. Average and best fitness functions.

4 Conclusions

In this study, an approach is proposed to perform optimization of backtest that based on indicators commonly used in technical analysis, on genetic algorithm. The parameters of the indicators used in technical analysis are chosen empirically or according to a certain rule. For example, since the parameters 12, 26 of the MACD indicator are applied in the Japanese markets, which are traded 6 days a week, they correspond to the number of 2-week and monthly bars. For algorithmic or manual transactions, these parameters vary for each market, financial asset and the monitored period. Traditionally, choosing the one that gives the best results on the historical data with the conventional approach can cause overfitting problems. In order to prevent this, verification with WFA and/or MCS after Backtest optimization creates a disadvantage. In this article, a genetic algorithm-based approach that converges to the best result is proposed, and in experimental studies, the proposed approach in all 5 different strategies outweighs conventional approaches.

Acknowledgments

This work was supported by the Scientific and Technological Research Council of Turkey (TUBITAK, Grant No. 121E733).

Declaration

The authors declare that the ethics committee approval is not required for this study.

References

[1] Christodoulaki, E., Kampouridis, M., & Kanellopoulos, P. (2022, May). Technical and sentiment analysis in financial forecasting with genetic programming. *In 2022 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CIFER)* (pp. 1-8). IEEE.

[2] Cheng, WK, Bea, KT, Leow, SMH, Chan, JYL, Hong, ZW, & Chen, YL (2022). A Review of Sentiment, Semantic and Event-Extraction-Based Approaches in Stock Forecasting. *Mathematics*, 10(14), 2437.

[3] Cheng, D., Yang, F., Xiang, S., & Liu, J. (2022). Financial time series forecasting with multi-modality graph neural network. *Pattern Recognition*, 121, 108218.

[4] Dai, C. (2022). A method of forecasting trade export volume based on back-propagation neural network. *Neural Computing and Applications*, 1-10.

[5] Mehtab, S., & Sen, J. (2022). Analysis and forecasting of financial time series using CNN and LSTM-based deep learning models. *In Advances in Distributed Computing and Machine Learning* (pp. 405-423). Springer, Singapore.

[6] Santur, Y. (2020). Deep learning based regression approach for algorithmic stock trading: A case study of the Bist30. *Gumushane University Journal of Science*, 10(4), 1195-1211.

[7] Klein, T. (2022). A note on GameStop, short squeezes, and autodidactic herding: An evolution in financial literacy? *Finance Research Letters*, 46, 102229.

[8] Sianturi, MS, & Kim, SS (2022). Technical Analysis and Value Investment in the Indonesia Stock Market. *Budapest International Research and Critics Institute (BIRCI-Journal): Humanities and Social Sciences*, 5(2).

[9] Achyutha, PN, Chaudhury, S., Bose, SC, Kler, R., Surve, J., & Kaliyaperumal, K. (2022). User Classification and Stock Market-Based Recommendation Engine Based on Machine Learning and Twitter Analysis. *Mathematical Problems in Engineering*, 2022.

[10] Shindler, M., Goodrich, MT, Gila, O., & Dillencourt, M. (2022). Beyond big O: teaching experimental algorithms. *Journal of Computing Sciences in Colleges*, 37(10), 23-36.

[11] Gao, R., Du, L., Yuen, KF, & Suganthan, PN (2021). Walk-forward empirical wavelet random vector functional link for time series forecasting. *Applied Soft Computing*, 108, 107450.

[12] Mukherjee, A., Singh, AK, Mallick, PK, & Samanta, SR (2022). Portfolio Optimization for US-Based Equity Instruments Using Monte-Carlo Simulation. *In Cognitive Informatics and Soft Computing (pp. 691-701)*. Springer, Singapore.

[13] Sang, B. (2021). Application of genetic algorithm and BP neural network in supply chain finance under information sharing. *Journal of Computational and Applied Mathematics*, 384, 113170.

[14] Chen, T. (2021). Evaluation method of development level of Science and technology finance in Heilongjiang Province based on genetic projection Pursuit model. *In 2021 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)* (pp. 805-809). IEEE.

[15] Wen, C., Yang, J., Gan, L., & Pan, Y. (2021). Big data driven Internet of Things for credit evaluation and early warning in finance. *Future Generation Computer Systems*, 124, 295-307.

[16] Wang, Q., Wang, R., & Ziegel, J. (2022). E-backtesting. *arXiv preprint arXiv:2209.00991*.

[17] Bailey, DH, & de Prado, ML (2021). How “backtest overfitting” in finance leads to false discoveries. *Significance*, 18(6), 22-25.

[18] Öntürk, A., Ulaş, M., Karabatak, M., & Santur, Y. (2022). Python for Finance: Open Source Tools and Development of an Algorithmic Robot. *3rd International Conference on Applied Engineering and Natural Sciences*

[19] Rodrigues, M. J. (2022). *Data Science for finance: automated investment recommendation with python* (Doctoral dissertation, Universidade Nova de Lisboa).

[20] Olivares, KG, Garza, F., Luo, D., Challú, C., Mergenthaler, M., & Dubrawski, A. (2022). HierarchicalForecast: A Reference Framework for Hierarchical Forecasting in Python. *arXiv preprint arXiv:2207.03517*.