WSB Journal of Business and Finance
Year 2024, Vol. 58, No.1
eISSN 2657-4950

Applying machine learning method in stock trading by indicator

Le Hoang Anh[a], Nguyen Van Nhat[b], Nguyen Luu Tuyen[c],

[a,b]*Ho Chi Minh University of Banking, 36 Ton That Dam street, District 1, Ho Chi Minh city, 700000, Vietnam*
[c]*Ho Chi Minh City University of Industry and Trade, 140 Le Trong Tan street, Tan Phu District, Ho Chi Minh city, 700000, Vietnam*

DOI 10.2478/WSBJBF-2024-0019

Abstract

The stock market is always considered a highly potential investment channel for the public. However, unpredictable fluctuations often require investors to continuously monitor and analyze the market with a huge amount of data. This research was conducted to implement machine learning methods in automated stock trading based on indicators, aiding investors in evaluating the effectiveness of trading strategies based on these indicators and suggesting the most appropriate investment portfolios, thus minimizing the time and effort spent on data processing. Specifically, the application process we developed and implemented involves four steps: (i) data collection, (ii) automated trading based on indicators (SMA, Bollinger Bands, RSI, MACD), (iii) building an optimal investment portfolio based on automated trading results using the Sharpe ratio method, and (iv) testing and evaluating the trading results with new data. Using data collected from VN30 stocks, the study results demonstrate that trading based on indicators and proposing an optimal investment portfolio yields high-profit rates and minimizes investor risks.

Keywords: SMA, Bollinger Bands, RSI, MACD, Sharpe ratio and optimal portfolio

1. Introduction

The stock market serves as a source of business capital and is an attractive, profitable investment channel for people. The stock market became a prominent and noteworthy point, especially when the COVID-19 epidemic occurred, leading to many areas of business activities being frozen. Stocks are highly profitable investment tools in the capital market, and the strong development of the stock market has attracted the attention of many investors who want to diversify their income and assets. According to information from the Ministry of Finance, in 2023, the number of new investor accounts increased by over 350,000 accounts, bringing the total number of accounts to more than 7.4 million accounts, equivalent to 7.5% of the population. (Binh Khanh, 2024). Along with the increase in new investors, businesses are also increasing their listing activities to access capital in the market. However, the attractive profits of stocks and the stock market come with fluctuations, unpredictability, and market risks. The stock market often reacts quickly to inflation, economic decline, and international geopolitical conflicts. Therefore, the stock market requires investors, when participating in the market, to master knowledge of economics, society, investment fields, finance, financial analysis, and the ability to update information continuously to evaluate and select appropriate stock codes and trading strategies as well as make accurate investment decisions.

Nowadays, the development of programming languages and artificial intelligence has unlocked many potential applications in assisting human decision-making across various fields (Chi & Chi, 2021; Jošić & Žmuk, 2022; Khan & Awasthi, 2019; Sumi et al., 2012; Yan, 2023). Among these, the application of technology in the finance and banking sector has become an appealing trend, attracting the attention of many scholars. The application of technology to stock trading is considered to create a big change in stock investment activities thanks to the ability to process huge amounts of data in a quick time (Cheng et al., 2022). In particular, programming languages have gradually become more popular and easier to learn, one of which is the Python programming language. Many investors or fund managers are using Python to collect, process data and calculate technical analysis such as indicators and investment performance indices (Cheng et al., 2022). Along with this development, many studies worldwide have applied programming languages to build automatic trading algorithms to help investors quickly make investment decisions (Cheng et al. , 2022; Frattini et al., 2022; Letteri, 2023). In Vietnam, research applying machine learning methods to recommend effective investment portfolios based on trading strategies is still limited. Most of these studies focus on stock price forecasting (Nhat & Trung, 2021; Truong Thi Thuy Duong, 2023; Tuyen, 2024; Kien & Tri, 2024). Therefore, in this study, the authors build an automatic trading strategy based on the indicator to select stocks with the highest profitability from the VN30 portfolio and then propose an optimal investment portfolio based on these stock codes to optimize investors' risk using the Sharpe ratio methodology. With this approach, we will build a stock selection process that ensures high returns and lowest risk for investors.

Section 2 presents the theoretical review of indicators, trading strategies based on indicators, and methods to optimize portfolio risk. Section 3 describes the research procedure and describes the data. Section 4 highlights the research results, and conclusions are presented in section 5.

2. Theoretical review of risk optimization indicators and methods

Technical indicators have become an effective tool for investors to grasp the price trend of stocks and then make buy/sell transactions. Indicators are mathematical functions integrated into technical analysis software and tested for good investment efficiency.

Simple Moving Average (SMA) is the most popular and easiest-to-understand financial analysis tool in technical analysis. The SMA is used to measure the price trend of a stock by averaging the closing price of the most recent price candle over a certain period of time (Hilpisch, 2020). One of the trading strategies based on the SMA line is that if the short-term SMA line crosses from below to the long-term SMA line, indicating that the market is increasing in price, a buying transaction should be made. And vice versa, if the SMA line crosses from above to the long-term SMA line, it shows that the market is bearish, so a sell transaction should be made. The authors use the above strategy with SMA20 and SMA50.

Bollinger Bands were created by John Bollinger in the 1980s and are used in financial markets to measure stock price movements and identify potential support and resistance levels. Bollinger Bands include the SMA line (can be adjusted according to strategy, usually 20-day SMA) and two upper/lower Bollinger lines with the formula SMA plus (for the upper Bollinger line) and minus (for the lower Bollinger line) with some standard deviations multiplied by the standard deviation of prices over a given period. The authors use a trading strategy based on signals from rebounding the band up or down for research. This rebound is because the upper and lower bands of the Bollinger bands act as dynamic support and resistance zones. Therefore, the Bollinger band trading strategy is specifically buying when the price exceeds the lower band and selling when the price exceeds the upper band.

Relative Strength Index (RSI) is used to determine the continuing trend of stocks, first introduced to the public in 1978 by mechanical engineer John Welles Wilder Jr. RSI compares the magnitude of recent price changes between stocks to determine whether a stock is overbought or oversold. The RSI indicator calculates the ratio between the average price increase and decrease in price over a certain period of time (usually 14 days) (Hilpisch, 2020). The value of RSI is expressed on a scale from 0 to 100. The specific RSI trading strategy is to buy when the RSI is below 30, indicating that the stock is oversold, and sell when the RSI is above 70, indicating that the stock is overbought.

Moving Average Convergence Divergence (MACD) was created in 1979 by inventor Gerald Appel. The MACD line reflects fluctuations and provides market buying and selling signals. MACD is the value found when subtracting the 26-day moving average (EMA) from the 12-day moving average (Hilpisch, 2020). One of the popular strategies using the MACD indicator is the trading strategy when MACD crosses the Zero line. Specifically, the strategy is a buy transaction when the MACD line crosses the Zero line from the bottom up or the MACD line shifts from negative to positive and a sell transaction when the MACD line crosses the Zero line from the top down or the MACD line shifts from positive to negative.

Sharpe Ratio measures how much profit is earned per unit of risk when investing in an asset or according to a business strategy. The Sharpe ratio is used to help investors understand an investment's return relative to its risk. Essentially, the Sharpe ratio is the average return earned in excess of the risk-free return per unit of risk. The greater the Sharpe ratio of the investment portfolio, the better the risk-adjusted performance. Thus, to optimize the risk of the investment portfolio, it is necessary to allocate the capital proportion of each asset in the portfolio so that the value of the Sharpe ratio reaches the maximum.

3. Research process

This study was conducted according to the following process:

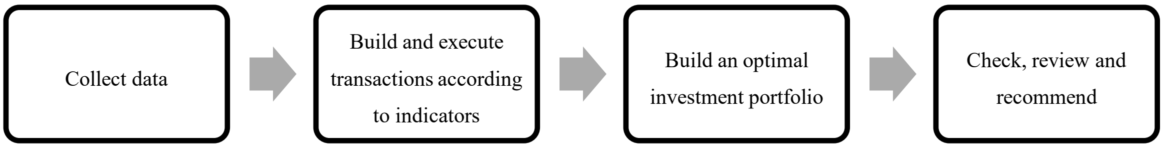

Fig. 1. Research process

The research team collected transaction price data within 5 years (2018 - 2023) of VN30 stock codes. Collected data is divided into two phases: the training phase (2018 - 2022) and the testing phase (2023). At the same time, the research team builds trading algorithms based on indicator trading strategies and portfolio risk optimization algorithms based on the Sharpe ratio method.

The research team trades stock codes based on indicator trading strategies during training. The 5 stocks with the highest return in each indicator strategy form an investment portfolio to maximize the portfolio's return. In addition, the research team optimizes risk for investment portfolios by calculating the capital allocation ratio to achieve the maximum Sharpe ratio or the highest portfolio risk-adjusted performance. The data used to optimize portfolio risk is data in the training phase. Then, trade the above portfolios along with the corresponding strategies in the test phase to find the most optimal portfolio and strategy based on the return rates of the investment portfolios.

4. Research results

4.1. Results of building investment portfolios:

After trading in the training phase, 4 investment portfolios are built corresponding to each indicator trading strategy as follows:

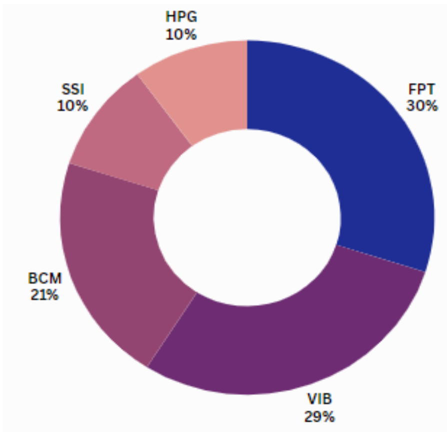

Fig. 2. Optimal investment portfolio for trading strategy according to the SMA indicator.

Based on Fig. 2, the optimal investment portfolio for the trading strategy according to the SMA indicator includes the stock codes FPT, VIB, BCM, SSI, and HPG with capital allocation proportions respectively of 30%, 29%, 21%, and 10% for the remaining codes. With this proportion, the maximum Sharpe ratio is 0.79, the portfolio's expected return is 19%/year and the estimated risk is 21%.

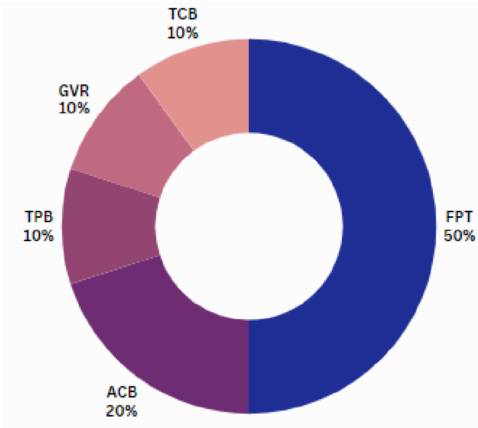

Fig. 3. Optimal portfolio for Bollinger bands trading strategy

Based on Fig. 3, the optimal investment portfolio for the Bollinger band trading strategy includes the stock codes FPT, ACB, TPB, GVR, TCB with capital allocation proportions respectively of 50%, 20% and 10% for the remaining codes. With this proportion, the maximum Sharpe ratio is 0.86, the portfolio's expected return is 18%/year and the estimated risk is 17%.

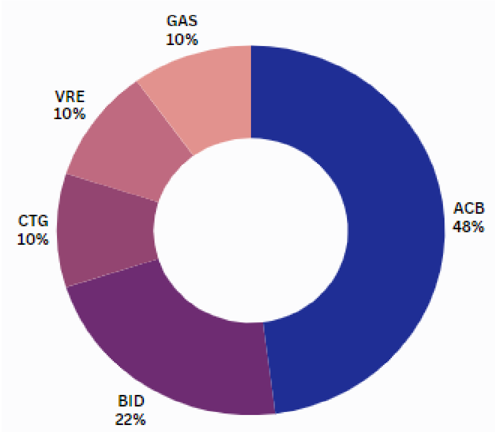

Fig. 4. Optimal investment portfolio for trading strategy based on the RSI indicator

Based on Fig. 4, the optimal investment portfolio for the RSI trading strategy includes the stock codes ACB, BID, CTG, VRE, and GAS, with capital allocation proportions respectively of 48%, 22%, and 10% for the remaining codes. With this proportion, the maximum Sharpe ratio is 0.28, the portfolio's expected return is 10%/year and the estimated risk is 24%.

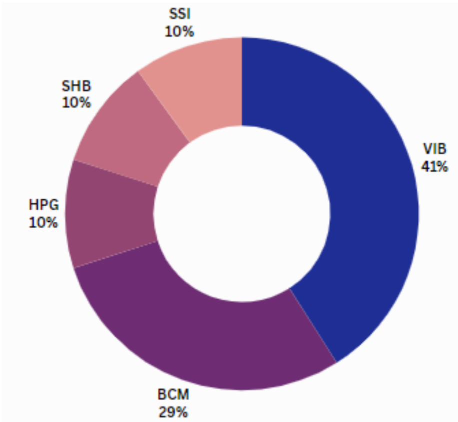

Fig. 5. Optimal investment portfolio for MACD indicator trading strategy

Based on Fig. 5, the optimal investment portfolio for the RSI trading strategy includes the stock codes VIB, BCM, HPG, SHB, and SSI, with capital allocation proportions of 41%, 29%, and 10% for the remaining codes.

With this proportion, the maximum Sharpe ratio is 0.68, the portfolio's expected return is 19%/year and the estimated risk is 24%.

4.2. Trading results of investment portfolios

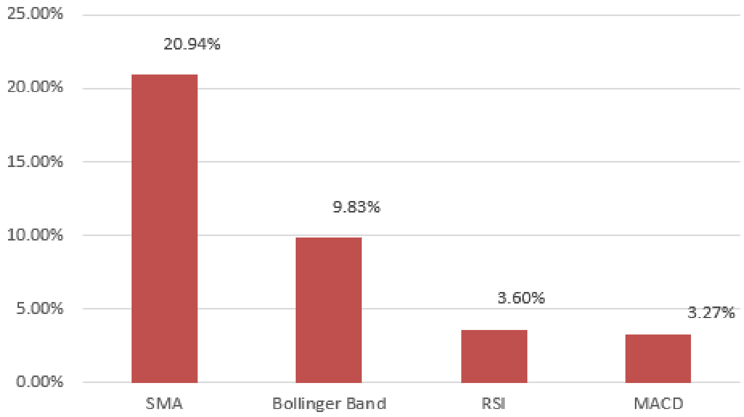

Fig. 6: Returns of portfolios by indicator after 2023 trading

Based on Fig. 6, it can be seen that all the investment portfolios, according to the indicator, are profitable. The investing portfolio, according to the SMA indicator, has the highest return of 20.94%, surpassing the expected return for 2023 of 19%. The final result of the process shows that the investment portfolio and trading strategy, according to the SMA indicator, are the most optimal for investors.

However, when actually trading, it is necessary to monitor and periodically check the investment portfolio's effectiveness (every 3 months, 6 months...) to make appropriate adjustments based on the most recent price data and allow investors to re-select investment strategies or change stocks in the portfolio.

5. Conclusion and recommendations

5.1. Conclusion

The authors used machine learning methods in this study to build optimal investment portfolios based on indicators. The results show that the research application process has practical value, allowing investors to select and test optimal indicator trading strategies and portfolios.

However, the research also has limitations, such as a limited number of stocks, indicators, and portfolio risk optimization methods. The research team hopes that future research can expand to the entire Vietnamese stock market, combine several indicators within the same trading strategy, and optimize portfolio risk using other methods.

5.2. Proposal for practical application

Based on the developed algorithm and the research results, we propose the following process for the practical application of machine learning methods in stock trading based on indicators:

Step 1: Collect data

Selecting stocks: Investors list stocks they are interested in according to their criteria, such as rankings, by industry...

Collect price data of selected stocks in the list:

- Data collected is transaction prices.
- Data collection period is at least 5 years.
- Data is divided into:
  - o At least 2 years of training data.
  - o Test data for at most 1 year.

Data is collected for 5 years from 2018 - 2023. Data is divided into training data (from 2018 to 2022) and testing data (2023).

Step 2: Build a trading strategy based on the indicators and optimize portfolio risk.

This study has built a trading algorithm based on indicators (SMA, Bollinger bands, RSI, and MACD), an algorithm to determine the optimal risk capital allocation ratio for the investment portfolios according to the Sharpe method. However, each indicator has different trading strategies and risk optimization methods. So, in this step, investors should select indicators, determine how the indicators are used to generate accurate buy and sell signals, and adjust the indicator's signal output thresholds to suit specific market conditions. Also, investors should build trading algorithms based on indicator trading strategies and risk optimization algorithms for investment portfolios.

Step 3: Trade and optimize the investment portfolios risk using training data

Warren Buffet once said that investment portfolios should be diversified but not too many; 3 to 6 assets per portfolio can be "all the diversification you need," and avoid investing in things they do not understand. Therefore, investors should choose 3 - 6 stock codes for the investment portfolio. In order to maximize profits, stocks are included in the investment portfolio based on the stock's return after trading using training data. The optimal risk capital allocation of portfolios is determined using training data.

Step 4: Trade optimal investment portfolio using test data

Execute trading transactions using test data with the same investment capital for the portfolios and capital allocation according to the corresponding optimal proportions. The results obtained after the trading transaction show investors which investment portfolio brings the highest profit corresponding to the most effective investment strategy.

Step 5: Implementation and actual monitoring

Based on the results in step 4, investors execute the investment portfolio and trade with the strategy according to the corresponding indicator. At the same time, investors monitor the portfolio's performance over time (3 months, 6 months, 1 year...) and adjust the portfolio when necessary to maintain the best performance by repeating the process.

REFERENCE

Binh Khanh. 2024. Soon put new technology systems into operation to ensure stock transactions. Retrieved from http://tuoitre.vn/som-dua-he-thong-cong-nghe-moi-vao-hoat-dong-de-dam-bao-giao-dich-chung-khoan-20240102084226245.htm

Cheng, C.-H., Tsai, M.-C., & Chang, C. 2022. A Time Series Model Based on Deep Learning and Integrated Indicator Selection Method for Forecasting Stock Prices and Evaluating Trading Profits. Systems, 10(6). doi: 10.3390/systems10060243.

Chi, Y. N., & Chi, O. (2021). Modeling and Forecasting of Monthly Global Price of Bananas Using Seasonal Arima and Multilayer Perceptron Neural Network. Econometrics, 25(3), 21–41. https://doi.org/10.15611/eada.2021.3.02

Frattini, A., Bianchini, I., Garzonio, A., & Mercuri, L. 2022. Financial Technical Indicator and Algorithmic Trading Strategy Based on Machine Learning and Alternative Data. Risks, 10(12). doi: 10.3390/risks10120225

Hilpisch, Y. 2020. Python for Algorithmic Trading: From Idea to Cloud Deployment 1st Edition. USA: O'Reilly Media.

Jošić, H., & Žmuk, B. (2022). A Machine Learning Approach to Forecast International Trade: The Case of Croatia. Business Systems Research Journal, 13(3), 144–160. https://doi.org/10.2478/bsrj-2022-0030

Khan, M. J. U. R., & Awasthi, A. (2019). Machine learning model development for predicting road transport GHG emissions in Canada. WSB Journal of Business and Finance, 53(2), 55–72. https://doi.org/10.2478/wsbjbf-2019-0022

Letteri, I. 2023. VolTS: A Volatility-based Trading System to forecast Stock Markets Trend using Statistics and Machine Learning (arXiv:2307.13422). arXiv. doi: 10.48550/arXiv.2307.13422

Nhat N. M., & Trung N. D. 2021. Quantitative trading application on Vietnam stock market. Asian Journal of Economics and Banking, 183.

Simanjuntak, I., Heriyant, H., Rochendi, A., Rahmawati, Y., Salamah, K., & Sulistiyono, S. 2023. Trading Simulation Using Python and Visualization on Streamlit with Machine Learning Decision Tree. Proceedings of the 2022 International Conference on Computer, Control, Informatics and Its Applications, 286–291. doi: 10.1145/3575882.3575937

Sumi, S. M., Zaman, M. F., & Hirose, H. (2012). A rainfall forecasting method using machine learning models and its application to the Fukuoka city case. International Journal of Applied Mathematics and Computer Science, 22(4), 841–854. https://doi.org/10.2478/v10006-012-0062-1

Truong Thi Thuy Duong. 2023. Forecasting the direction of stock index fluctuations using enhanced algorithms. Retrieved from http://hvnh.edu.vn//tapchi/vi/thang-5-2023/du-bao-chieu-bien-dong-cua-chi-so-chung-khoan-bang-thuat-toan-tang-cuong-truong-thi-thuy-duong-10755.html

Tuyen T.D. 2024. Evaluating the performance of the LSTM-GRU complex model: Case study on forecasting an index measuring stock price fluctuation trends on the Ho Chi Minh stock exchange. Can Tho University Science Magazine, 60(1). doi: 10.22144/ctujos.2023.232

Trung Kien Dang et al 2024. Factors affecting the proftability of food companies listed on the vietnam stock market, 17(1), 70 - 88, Electronic Journal of Applied Statistical Analysis. DOI: 10.1285/i20705948v17n1p69

Yan, N. (2023). Construction of International Trade and Investment Platform Based on Artificial Intelligence Technology. Applied Mathematics and Nonlinear Sciences, 9(1). https://doi.org/10.2478/amns.2023.2.01694