International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Super Stock Trading: Automation in Reinforcement Learning with Advanced Multi-Indicator Confirmations

Aditya Sharma[1], Arup Kadia[2*], Suryansh Kumar[3], Rajraushan Kumar[4]

[1,3,4] Student, BCA (DS & AI), Faculty of Information Technology & Engineering, Gopal Narayan Singh University, India
[2] Assistant Professor, Faculty of Information Technology & Engineering, Gopal Narayan Singh University, India

Abstract
One of the main reasons that stock market trading can be very difficult is the uncertainty that comes with price fluctuations as well as the noise from the price movements, which together often mislead traders when they use technical indicators that are not complemented by others. The proposed research work presents an algorithmic stock trading framework based on Reinforcement Learning (RL) that links the use of Simple Moving Average Crossover (SMAC) for trend identification, along with Average Traded Volume (ATV) and Put-Call Ratio (PCR) for signals confirmations sentiment, respectively. SMAC is the main source of buy and sell signals, while ATV determines the strength of the market, and PCR shows the sentiment from the derivatives market. The framework is built on Proximal Policy Optimization (PPO) which learns trading policies by adapting them from historical market data. The model looks at trend direction, volume dynamics, sentiment conditions, and current position status in order to decide whether to buy, sell, or hold. Instead, a risk-aware reward formulation that incorporates transaction costs and drawdown penalties is implemented to ensure a stable performance. The results of the experiments show that the SMAC-ATV-PCR-PPO strategy suggested by the authors is superior to the traditional SMAC-based trading methods in terms of risk-adjusted returns and also in terms of decreasing to a significant extent the occurrence of false signals and their handling.

Keywords: Reinforcement Learning, Algo trading, Market Sentiment, Proximal Policy Optimization (PPO), Simple Moving Average Crossover, Average Traded Volume, Put Call Ratio

1. Introduction

The financial markets are by their very nature intricate, ever-changing, and exceedingly uncertain. This has made the task of achieving steady and trustworthy trading a hard one. Price fluctuations are the result of a mixture of macroeconomic factors, the behavior of investors, liquidity in the market, and the sentiment among the participants, leading to the formation of patterns that cannot be easily described or predicted. The use of traditional rule-based trading strategies is often limited due to their simplicity and interpretability. They tend to be less flexible in meeting the rapidly changing market conditions and they also generate false signals, especially in times of heightened or stagnant market activities. It is therefore no surprise that the use of intelligent and adaptive trading systems that can learn the best decision-making strategies right from the market data has started to gain momentum [1].

RL recently, become a predominant automation model in algorithmic stock trading due to its capability to address sequential time series-based decisions and its optimization of long-term trading performance objectives and accuracy. Supervised learning that is based on labels fixed in advance is the approach opposite to RL.[2] The RL agents get their training and thus develop their trading policies through ongoing interaction with the environment and getting rewards that are cumulative throughout the whole period. Several studies have indicated the success of RL in trading and portfolio optimization where it has been found that the profitability and risk-adjusted returns have been better than those of the traditional. Moreover, according to survey-based studies, financial markets are said to be very adaptable to dynamic and uncertain environments and thus RL-based frameworks are particularly suitable for them [3].

Structured features through technical indicators are an indisputable benefit that produces better-performing models due to the ability to create more interpretable structures. Training a "pure" reinforcement learning ("RL") model using only the price data results in several disadvantages, including: slow convergence rates; difficulty achieving stable learning; & lack of interpretability. The most recent research shows that using technical indicators (e.g., moving average, momentum oscillator, and other types of indicators) in conjunction with an RL-based model provides faster convergence rates, more reliable learning results, and improved interpretability. The results from these studies indicate that noisiness in the data can be effectively reduced and policy strength can be increased when the domain knowledge is combined with RL [4].

The Simple Moving Average (SMA) crossover is perhaps the most nonchalant and simple technical indicator among the different indicators to identify trends. SMA uses the price data available for the past few days and averages them for long and short periods, and thus gives buy/sell signals when the two periods' average prices cross over. However, several researchers have criticized the SMA methods for being too simple, as for example they always tell the trader to enter or exit after a significant movement has already taken place [5]. As a result, the researchers have turned their attention toward trying to propel SMA's progress through the implementation of different approaches including the applying machine learning and neural networks for dynamically changing parameters.

Indicators of volume and sentiment, on the other hand, have been recognized as the essential adjuncts to price-based signals in trading decision making. Average Traded Volume (ATV) is a measure of the strength of market participation and is also used as a tool for validating the credibility of price movements. Research asserts that traders using the volume information along with the normal analysis of price, signals, and so on can filter out weak or low-confident signals with the consequence of doubling their trading performance [6]. Likewise, derivatives-based sentiment measures like the Put-Call Ratio (PCR) also point to such situations, which provide the market, behaviors, and even investors' positioning insights, thus, granting yet another price trend confirmation layer. There have been studies focusing on money flow and sentiment dynamics which have provided evidence of the power that co-occurrence of price, volume, and sentiment indicators has in stronger market interpretation [7].

As a result of these findings, the present research suggests an innovative trading framework based on hybrid reinforcement learning, which would combine Simple Moving Average Crossover (SMAC) for trend detection and Average Traded Volume (ATV) and Put-Call Ratio (PCR) as confirmation indicators. The innovative framework has Proximal Policy Optimization (PPO) as its core for the purpose of acquiring adaptive trading policies that are balanced in terms of both the profit and the risk [8]. By taking into consideration trend direction, market participation, and sentiment conditions jointly, the proposed strategy will eliminate false signals, reduce over-trading, and thus, enhance return on the risk-adjusted basis. This work adds to the existing body of work by showing that informative technical indicators can indeed be effectively used along with reinforcement learning for building robust and market-ready algorithmic trading systems.

2. Literature review

The nature of algorithmic trading has greatly impacted the financial analysis of the markets and one of the main reasons for that is the use of machine learning and reinforcement learning (RL) in particular. Different types of traders, both retail and institutional, have been relying on technical analysis tools like moving averages and volume indicators for many years. However, a recent trend is noted in the literature where there is a shift in practice toward the use of hybrid systems that integrate different financial indicators with adaptive learning algorithms. This literature review focuses on progress made in building reinforcement learning based (RLB) trading systems, particularly SMA crossover strategies, volume indicators, and sentiment-based indicators integrated into intelligent trading systems.

Reinforcement Learning is the most appropriate method to handle sequential decision-making problems in changing environments like the financial markets. Supervised learning models use labelled data, while RL Agents will discover optimal trading strategies through trial-and-error with the market, and subsequently maximise their rewards. Numerous papers have presented the use of RL for trading and portfolio management positively and successfully. Li et al. (2021) extensively reviewed deep reinforcement learning-based trading systems and acknowledged the RL models' capability to deal with the changing and nonlinear nature of market behavior. In the same vein, Kumar and Singh (2021) pointed out that policy-gradient and actor-critic techniques provide better adaptability in algorithmic trading setups. The use of RL in stock trading has been validated by recent empirical studies. Xiong et al. (2023) have utilised deep Q-learning in automated stock trading and the outcome was better than that of traditional rule-based strategies, especially when technical indicators were incorporated in the learning process. Yu et al. (2022) highlighted the importance of feature selection and reward design for reinforcement learning in financial markets to get reliable and steady trading performance through a review of the recent trends and obstacles facing this application. Huang et al. (2023) were talking about model-free RL algorithms being the best in comparison to heuristic strategies when accompanied by informative state representations during times of market volatility.

The Simple Moving Average (SMA) crossover strategy is still one of the most commonly used tools in technical analysis, mainly because of its simple application and easy understanding. The signals to buy and sell are produced by the crossing of the short-term and long-term moving averages indicating the possibility of trend reversals (Batten et al., 2014). Nevertheless, the SMA crossover strategies have been very popular, but their main drawback has been lagging effects, and they also produce frequent false signals especially when the market is sideways or very volatile. One way to deal with these drawbacks is through the use of adaptive and data-driven approaches. Jha et al. (2025) for instance, showed that the combination of optimization techniques with machine learning could lead to financial modeling accuracy that is significantly higher. This serves as a hint that static SMA parameters may not be adequate in dynamic markets. Kadia et al. (2025) made a comparison between traditional SMA-based strategies and adaptive learning-enhanced models, and they found the former to outperform the latter in trending markets. All these have further established the fact that rather than being phased out, SMA-based strategies need either confirmation mechanisms or to be placed in adaptive frameworks in order to be effective in the trading environments of today.

The trading volume is a big factor in confirming or denying price changes and in judging the level of market participation. Average Traded Volume (ATV) is the measure that most often determines the power and longevity of price trends. Financial theory claims that volume is a reflection of the new information flow, hence it becomes a vital factor for the market interpreting. The use of the volume information in trading models adds to the context that price-based indicators alone cannot provide. Dey et al. (2024) noted that trading volume inclusion in a machine learning model significantly raised stock price prediction accuracy. Their research indicated that the volume surges during the SMA crossovers indicate strong and thus reliable trends. Moreover, Adhikary, S., et al. (2025) conducted a similar study by a volume-weighted RL-based trading strategy and found that validating trade signals with volume reduced overtrading and hence improved the total returns. Such studies strongly suggest that volume confirmation is a very good method for filtering out weak or fake signals.

Market sentiment has gradually gained the recognition of being a major factor in trading decision-making. One of the most popular sentiment indicators reflecting the expectations and positioning of the investors is the Put–Call Ratio (PCR), which is based on the data from the derivatives market. Generally, a high PCR shows that investors are pessimistic about the market, while a low PCR is an indicator of optimism. The utilization of sentiment indicators such as PCR in trading models allows the models to consider the psychological and behavioral factors in the market which are not directly revealed by the price data. Yu et al. (2022) stressed that it was necessary to combine sentiment and momentum indicators in reinforcement learning models to gain a better understanding of market dynamics. Likewise, Wu and He (2023) proved that the integration of technical indicators with sentiment-aware reinforcement learning frameworks increases the accuracy of trading and cuts the losses down. The evidence from these studies points out that sentiment-based indicators like PCR can be utilized effectively as confirmation tools when used together with trend and volume signals. The reinforcement learning approaches that make use of multiple technical indicators are increasingly backed by recent literature. Hybrid models are applicable when the agent's decision space is represented by structured features like SMA signals, volume thresholds, and sentiment indicators instead of just raw price data. One of the early studies by Zhang et al. (2020) successfully combining SMA crossover strategies with reinforcement learning showed that the timing of entry and exit was much better than that of traditional threshold-based methods. Furthermore, Li et al. (2021) discovered that the input environments based on multiple indicators lead to quicker learning and better risk-adjusted performance in deep reinforcement learning models.

Price-only RL agents were compared to those using technical indicators to enhance their performance in the study by Huang et al. (2023). The latter type of agents consistently scored higher than the corresponding baseline models, especially in the case of high volatility periods. Kadia et al. (2025)

bolstered these conclusions by introducing volume confirmation along with technical indicators in an RL framework, thus attaining better trading precision and lower drawdowns. The cumulative impact of these studies is that the hybrid indicator–RL models are effective.

2.1. Research Gaps and Opportunities

Although there has been considerable development, the literature still has several gaps remaining. To start with, the majority of the trading systems that are based on reinforcement learning take advantage of deep learning architectures that are quite complex and their lack of interpretability limits their practical usage. The RL models that are not only interpretable but also based on well-established trading principles are in demand. Another point is that, while SMA, volume, and sentiment indicators have been the subject of individual studies, very few researchers have looked at their simultaneous integration in a unified reinforcement learning framework. The focus of many studies has been on short-term or limited datasets, neglecting to take into account transaction costs and the long-term perspective. The present study is based on these gaps and suggests a hybrid trading framework that combines Simple Moving Average Crossover, Average Traded Volume, and Put–Call Ratio with a Proximal Policy Optimization (PPO)–based reinforcement learning model. The fusion of trend detection, validation of liquidity, and confirmation of sentiment is the way, that this research intends, to produce a trading system that is more reliable, interpretable, and adaptive to the real-world market conditions.

3. Methodology and model specifications

In this research paper we have created a state-of-the-art Multi-Indicator Trading System which combines historical technical indicator data along with Reinforcement Learning Model to give traders more accurate and reliable stock market trade decisions. By combining the use of Simple Moving Average Crossover (SMAC) in order to identify trends with Average Traded Volume (ATV) and Put-Call Ratio (PCR) to validate and filter out undesirable trades, this technique has a number of advantages for the money managers of today. A reinforcement learning agent based on Proximal Policy Optimization (PPO) is created to make trade execution even better by absorbing lessons from the past behavior of the market.

3.1. Data Acquisition

The stock market numbers are gained from trustworthy and commonly accepted financial sources. The information can be gathered from leading stock markets like the National Stock Exchange (NSE) and Bombay Stock Exchange (BSE) in India, the New York Stock Exchange (NYSE) in the US, and other worldwide exchanges. The dataset includes Open, High, Low, Close, and Volume (OHLCV) data as well as the necessary derivatives data for calculating the Put–Call Ratio.

The historical data are utilized to train and evaluate the reinforcement learning model, while the real-time data are used for signal generation in live market conditions. In this research, the OHLCV data for a five-year period of selected large-cap stocks on the NSE are considered.

3.2. Data Processing

Financial time-series data can be incomplete because of market off days, or temporary stoppages of trading. Missing values are filled through the forward-fill method to keep the continuity. The dataset is from the beginning to the end to keep the time order which is necessary for the time-series learning process.

The unnecessary attributes are eliminated, and only the necessary features for indicator calculation and model training are kept.

3.3. Calculation of Simple Moving Average Crossover (SMAC)

To identify trends, the Simple Moving Average (SMA) is used as the main indicator. By calculating the mean price over a certain time, the SMA reduces the impact of short-term price swings and reveals the direction of the market.

$$SMA(t,n) = \frac{1}{n}\sum_{i=0}^{n-1} P_{t-i} \qquad 1$$

where Pt is the closing price at time t, and n is the window length.

Two SMA values are computed:

- Short-term SMA
- Long-term SMA

Trading signals are generated as:

Golden Cross (Buy Signal) generated when Short-term SMA crosses above long-term SMA and **Death Cross (Sell Signal)** are generated when Short-term SMA crosses below long-term SMA

Although SMAC is effective for trend detection, it often generates false signals in sideways or low-volume market conditions.

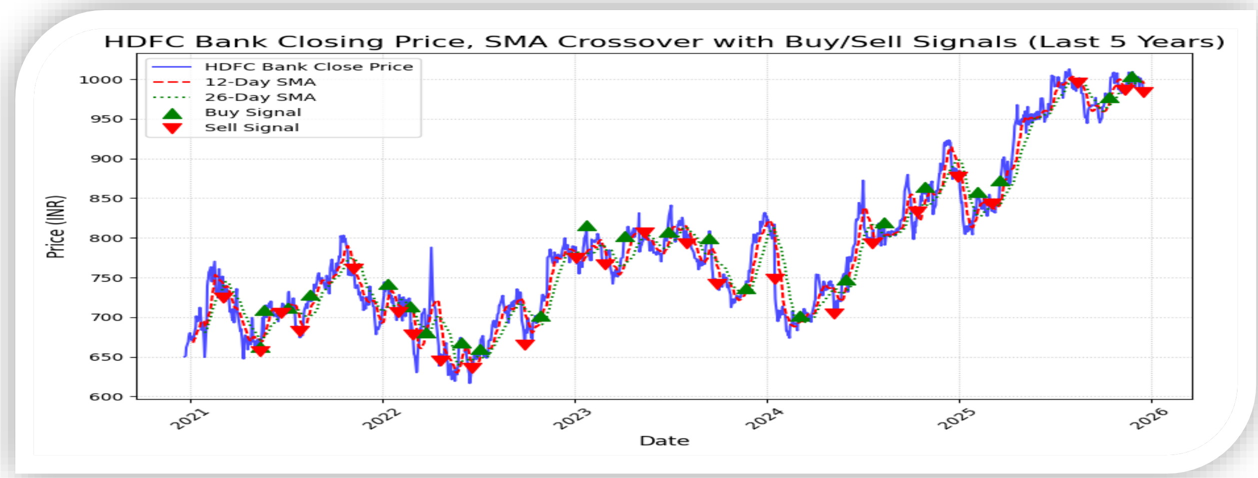

Figure 1: 12 Days and 26 Days SMA Crossover of HDFC Bank

Figure 1 illustrates the 12-day and 26-day moving averages and their crossovers over the past 5 years. The orange color represents the 12-day SMA while the green color illustrates the 26-day SMA. A golden crossover occurs when the orange line intersects the green line from below to above and a buy signal is generated thus represented by the green triangle, while the red triangle indicates the death crossover or sell signal which occurs when the orange line crosses the green line from above to below.

3.4. Average Traded Volume (ATV) Confirmation

For reinforcement of SMAC signals, Average Traded Volume (ATV) was applied as a volume confirmation indicator. ATV denotes the average volume of shares traded during a specified time and shows the strength of market participation.

$$ATV_t = \frac{1}{m}\sum_{i=0}^{m-1} V_{t-i} \qquad 2$$

where Vt represents trading volume at time t.

A SMAC signal is considered valid only if:

$$V_i > ATV_t \qquad 3$$

Figure 2: MFI indicator of Reliance Industries (1st June,2020 – 31st May,2025)
Figure 2 represents money flow indicator. Two threshold line 20 and 80 lines represents oversold and over brought respectively. Below 20 confirms buy signal generated by SMA and above 80 confirms sell signal generated by SMA.

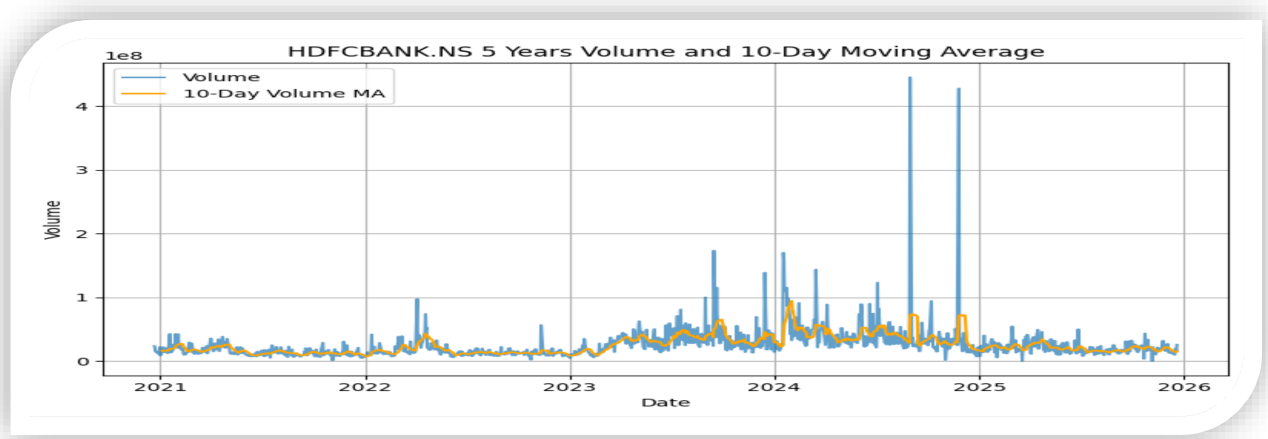

Figure 2: Average Traded Volume 10 Days MA in HDFC Bank

Concerning this caveat, price movements are purported to have significant backing in the form of trading activity.

3.5. Put–Call Ratio (PCR) Sentiment Analysis

The Put–Call Ratio (PCR) is used to measure the mood of the market by analyzing the data from the derivatives market. The ratio shows the relationship between the total open interest of put options and call options.

$$PCR = \frac{\text{Total Put Open Interest}}{Total\ Call\ Open\ Interest} \quad 4$$

- PCR down direction bearish sentiment
- PCR up direction bullish sentiment

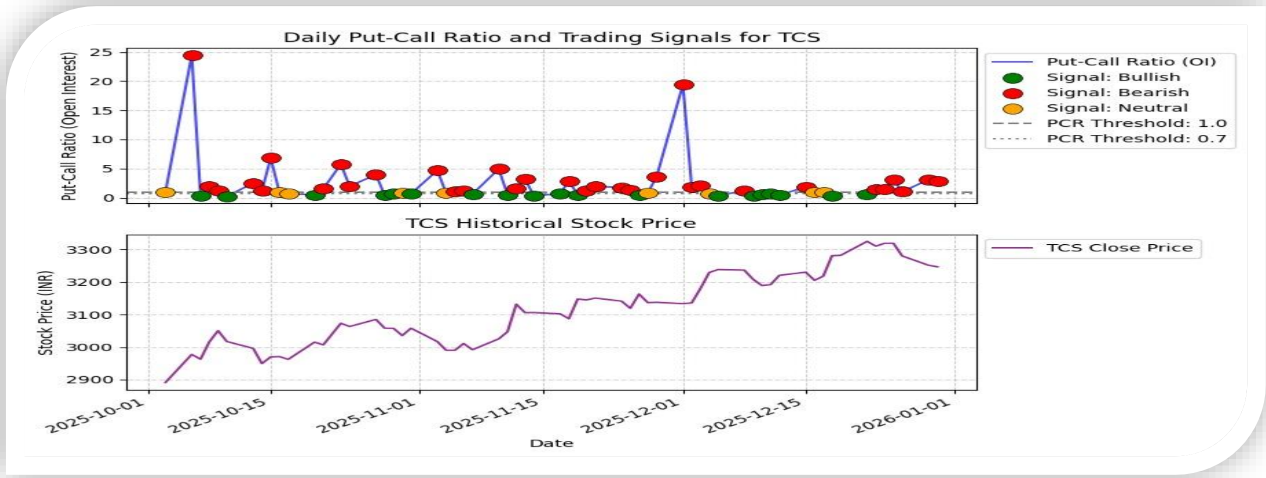

Figure 3: PCR-based trading strategy in TCS

PCR confirms directional signals and discourages you from going against the market's mood.

3.6. Reinforcement Learning Using PPO

To enhance the trade execution, a reinforcement learning agent that uses the Proximal Policy Optimization (PPO) method is applied.

- **State Space:** SMAC signal, ATV confirmation, PCR value, and position status
- **Action Space:** Buy, Sell, Hold
- **Reward Function:** Net profit adjusted for transaction costs and penalties for incorrect trades

The PPO objective function is given by:

$$L^{PPO}(\theta) = E[\min(r_t(\theta)A_t, clip(r_t(\theta), 1-\in, 1+\in)A_t] \quad 5$$

where rt(θ) is the probability ratio, At is the advantage function, and ϵ is the clipping parameter.

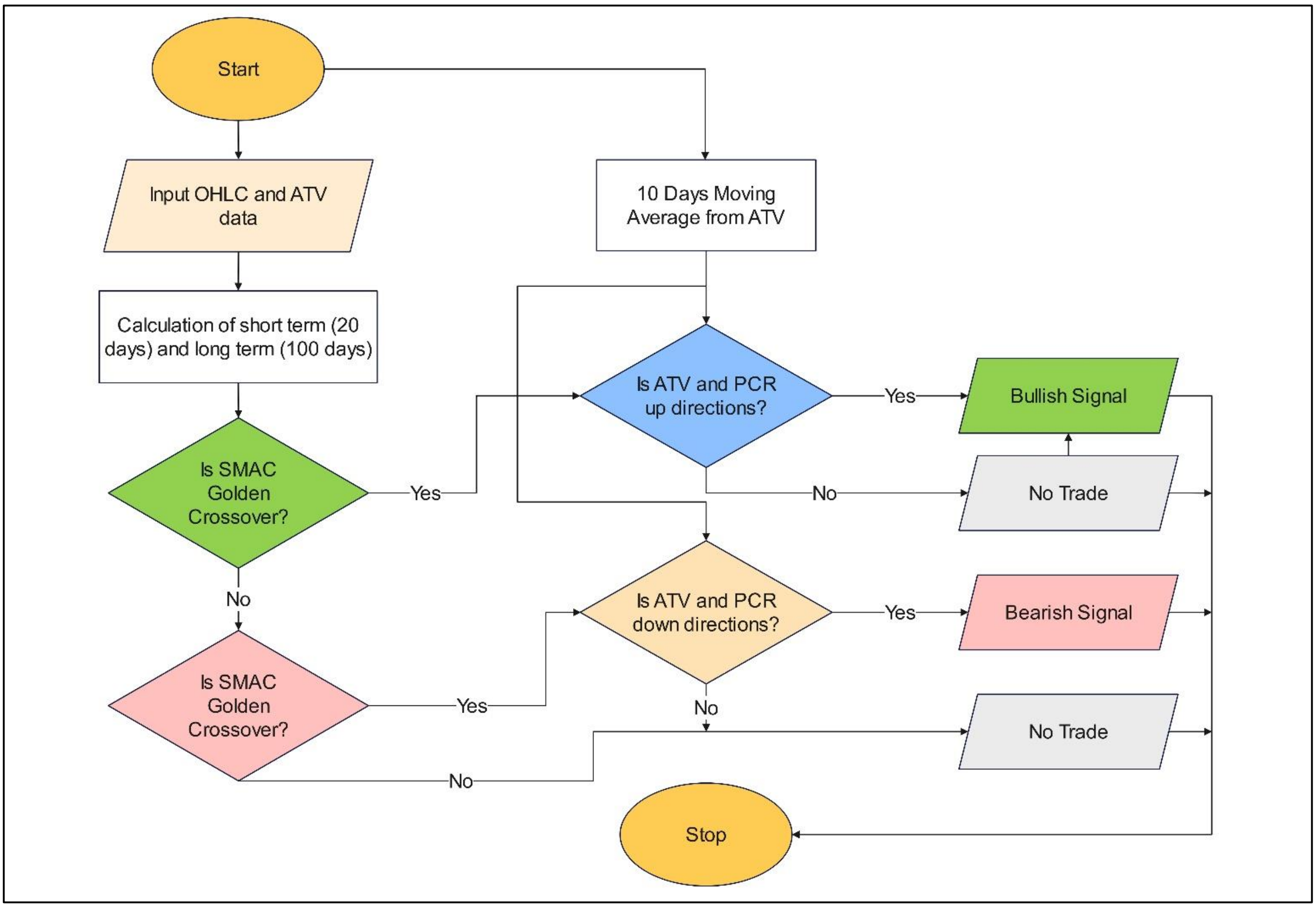

Figure 4: Flowchart of the proposed SMAC–ATV–PCR trading strategy

3.7. Summary of Methodology

Initially, the Simple Moving Average Crossover (SMAC) is set as the main mechanism for detecting the market trends, as well as for the creation of initial buy and sell signals. Moreover, an additional confirmation indicator is added to overcome the problem of false signals frequently generated by SMAC during the periods of low volume or sideways movement in the market. Average Traded Volume (ATV) is the indicator confirming the strength of price movements; thus, it makes sure that trade signals are backed up by the appropriate market participation. Furthermore, as a sentiment indicator based on the derivatives market, Put–Call Ratio (PCR) is deployed to estimate bullish or bearish market expectations and thereby filter out trades that go against the current sentiment. To execute the trades, the use of a reinforcement learning agent based on Proximal Policy Optimization (PPO) is implemented to make the best possible trade execution decisions. The agent learns from past interactions with the market, thus it adjusts its actions of buy, sell or hold during the course of the trading to make the highest returns while applying risk management. In conclusion, the suggested approach innovatively unites trend, volume, and sentiment data in an adaptive learning framework leading to a very strong and profitable choice of an automated trading system.

4. Empirical Result

The proposed trading framework has undergone an extensive backtesting process with the use of historical stock market data for the purpose of performance evaluation. SMAC's strategies versus SMAC's

International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

performance. KPI's measure the success of SMAC strategies. SMAC strategy performance is measured using KPI'S return on investment (ROI), Average profit/loss per-trade ($-totals), Trade Accuracy (% profitable trades to total trades completed, Sharpe ratio (risk-adjusted return on investment) and Sortino Ratio (risk-adjusted return on investment).
An increase in either KPIs , ROI or Average profit/loss per-trade($-totals), implies a decrease in both Sharpe and Sortino Ratios and vice versa.

The experimental evaluation includes a selection of ten large-cap stocks which are part of the Nifty 50 index and those are Reliance Industries, HDFC Bank, TCS, Infosys, ITC, ICICI Bank, Sun Pharma, Hindustan Unilever, Power Grid, and Asian Paints. In order to ensure robustness across different trading horizons, backtesting is conducted using 5-minute, hourly, and daily price data which represent short-term, medium-term, and long-term trading scenarios respectively.

4.1. Short-Term Performance (5-Minute Time Frame)

The basic SMAC-only strategy in the short-term analysis of 5-minute price data over a month generates a huge trade volume. Even though this is escalating market participation, it is also the reason for trading too much, getting a high number of bad signals, thus making the overall profit and increasing the risk of trading.

Average Traded Volume (ATV), when used as a confirmation mechanism, almost doubles the performance. ATV confirms SMAC signals only during the times of market participation that is sufficient. Consequently, non-profitable trades are lessened, and trade precision is increased. Additionally, by using Put–Call Ratio (PCR), the trading decisions are taken in line with derivatives-market sentiment, thus making it possible to keep away from trades that are against the prevailing market expectations. Henceforth, the SMAC–ATV–PCR strategy reveals the accumulation of profit and trade accuracy that is superior when compared to the SMAC-only tactics.

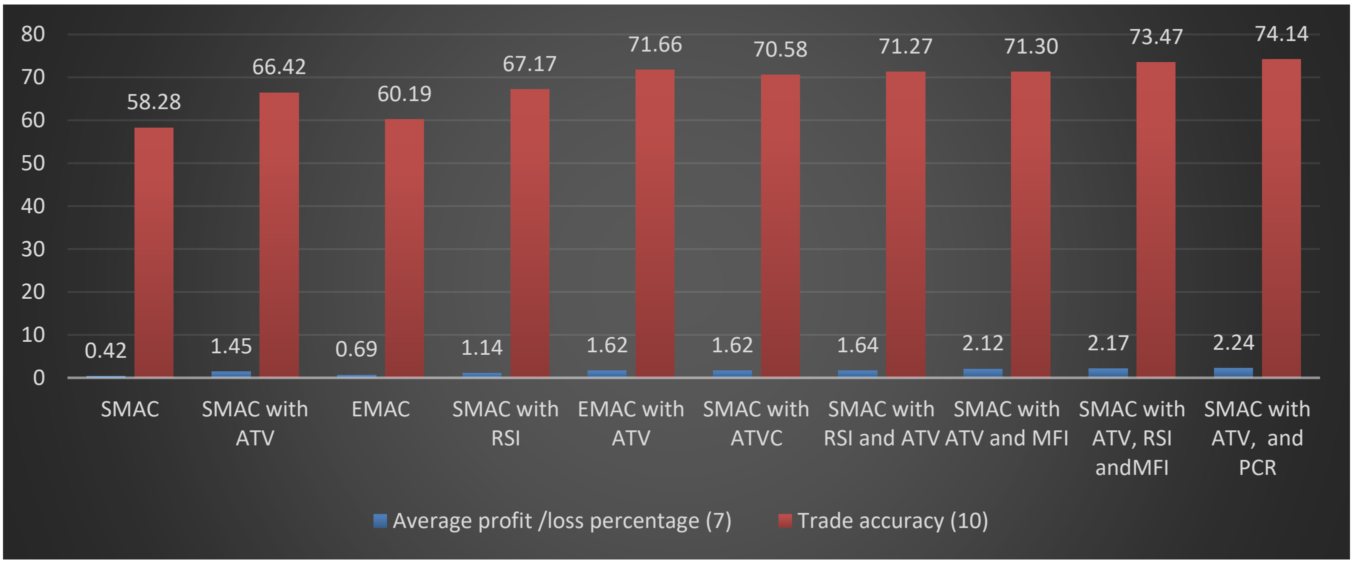

Figure 4: Cumulative profit and Trade accuracy are compared in 10 strategies in 1 month 5-min time frame

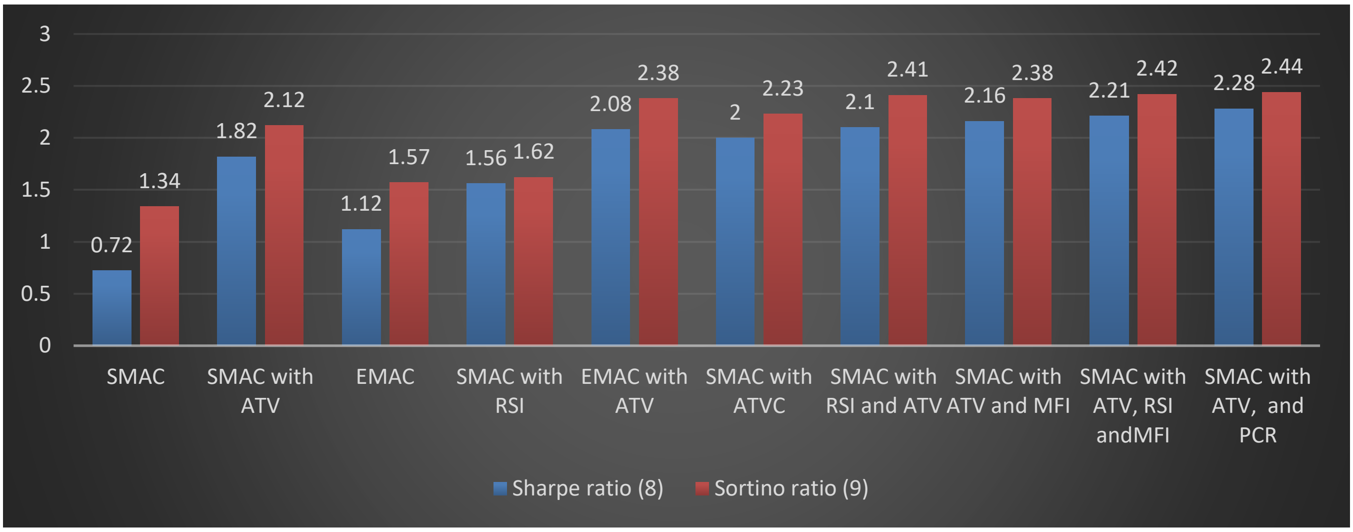

Figure 5: Sharpe ratio and Sortino ratio are compared in 10 strategies over 1 month 5-min time frame

4.2. Medium-Term Performance (Hourly Time Frame)

The medium-term evaluation with hourly data for a one-year period shows even more the power of confirmation through multi-indicators. The strategy of using SMAC only is suffering continuously from the range-bound and low-volume market phases that produce frequent false signals. On the other hand, the use of ATV limits overtrading by allowing trades to take place only in strong participation periods.

With the introduction of PCR, decision-making is even more improved through the addition of market sentiment coming from the derivatives market. It makes buy and sell decisions get in line with the overall market expectations. When such a combination is made with PPO-based reinforcement learning, the strategy shows a rise in the execution efficiency which is leading to a higher accuracy in trade and more stable profit growth. The Sharpe and Sortino ratios also go up which means better risk-adjusted performance.

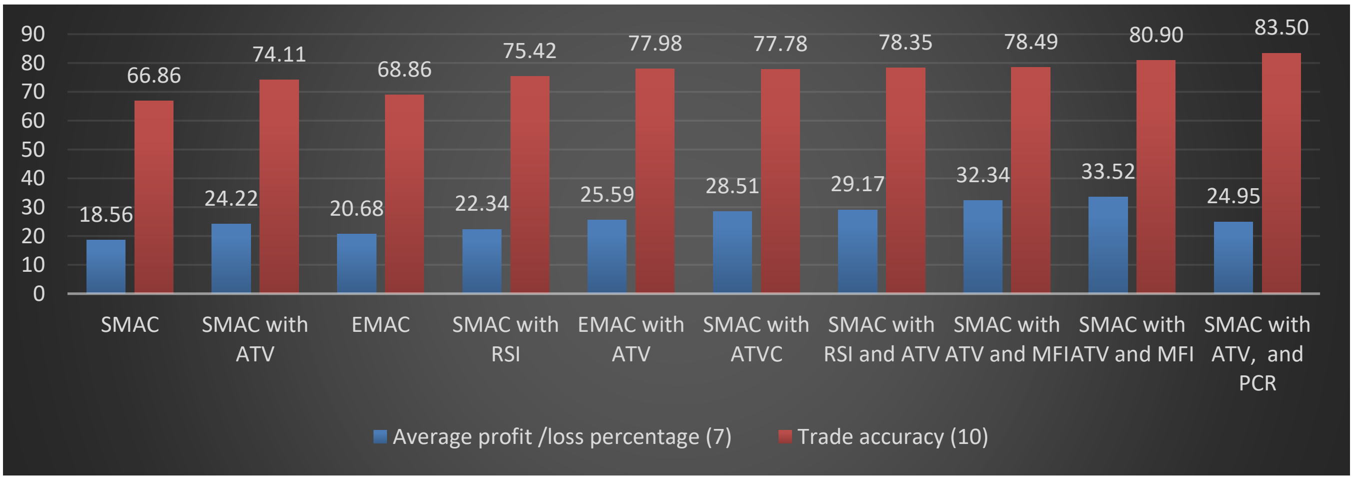

Figure 6: Cumulative profit and Trade accuracy are compared in 10 strategies in 1-year's hourly time frame

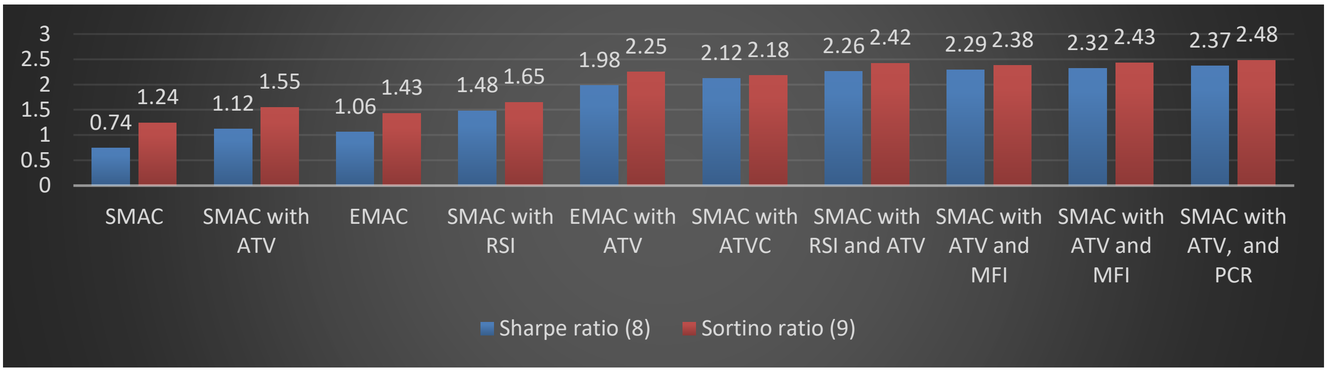

Figure 7: Sharpe ratio and Sortino ratio are compared in 10 strategies over 1-year hourly time frame

4.3. Long-Term Performance (Five-Year Period)

The long-term analysis that was done over a five-year period gives proof of the strength and reliability of the suggested framework. The SMAC-only strategy is characterized by performance instability because of the prolonged sideways markets and sudden volatility changes. The use of ATV assures that the trend signals are backed by strong market participation, which is an effective way of eliminating false breakouts.

Performance is even more improved by the inclusion of Put-Call Ratio (PCR) as it brings in sentiment information from the derivatives market. PCR acts as a guide to the traders by showing the mental state of the whole market, thus preventing them from making mistakes that are caused by misinterpretations of the sentiment. The SMAC-ATV-PCR strategy together with PPO has the highest cumulative returns, increased trade precision, and the best risk-adjusted performance over all tested time horizons.

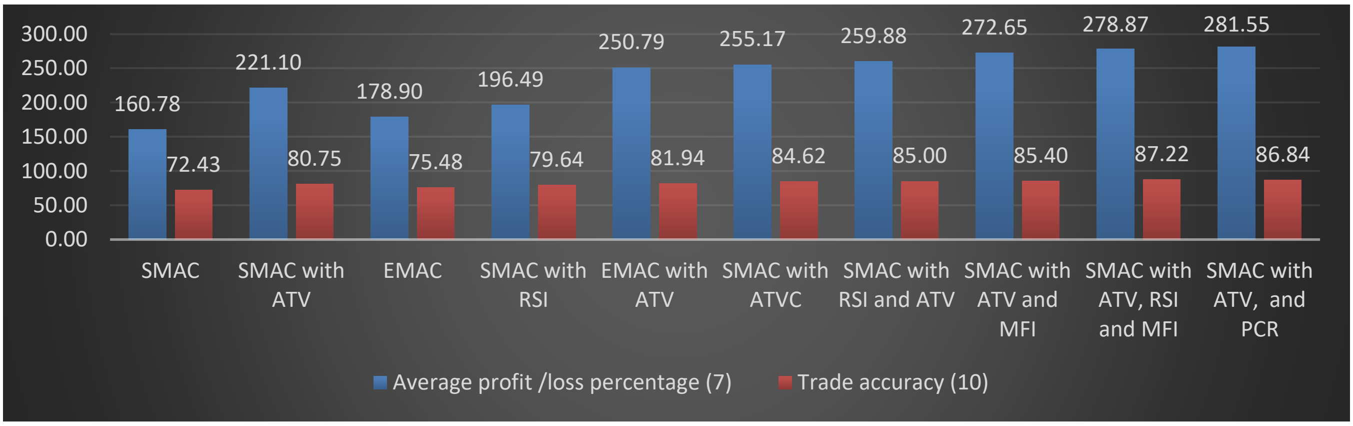

Figure 8: Cumulative profit and Trade accuracy are compared in 10 strategies in 5-years

International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

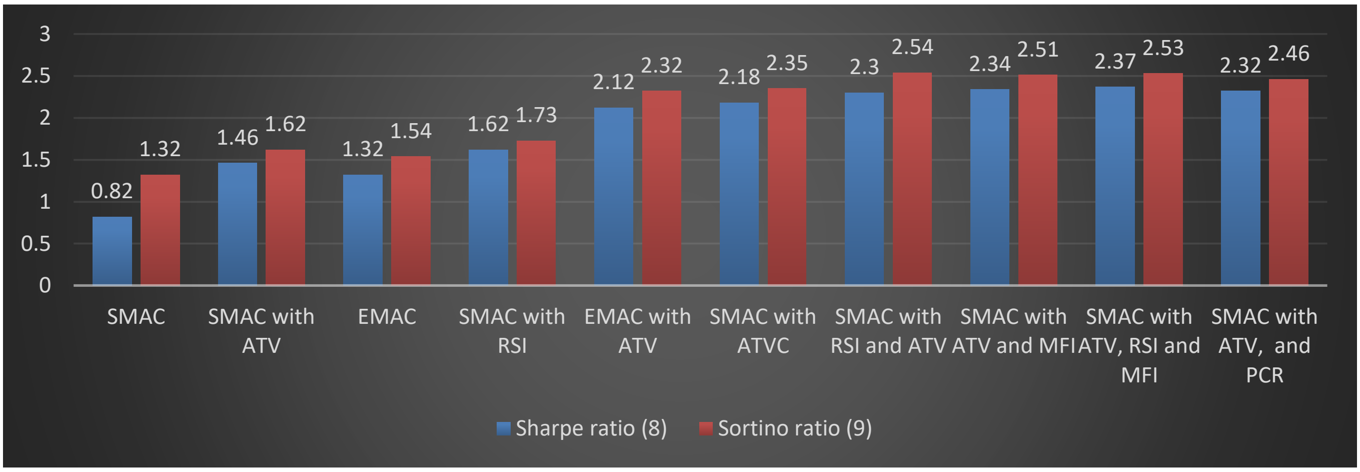

Figure 9: Sharpe ratio and Sortino ratio are compared in 10 strategies over 5-years

4.4 Overall Result Interpretation

Single indicators alone cannot filter the noise in the market or indicate when false signals have occurred, as demonstrated by the results of this analysis. ATV and trend confirmation confirm the existence of a trend, while PCR supports trend validity through the use of market sentiment. A model that incorporates PPO (Policy Payouts) and reinforcement learning can learn from past performance how to execute trades in the best manner possible, including what the best time to buy, sell, and hold is.

In general, the SMAC–ATV–PCR framework not only consistently provides greater profitability but also higher trade precision and better risk-adjusted returns, thus making it an excellent choice for automated trading in real life.

5. Conclusion

The paper has introduced a smart multi-indicator stock trading framework that brings together Simple Moving Average Crossover (SMAC) with Average Traded Volume (ATV) and Put–Call Ratio (PCR), all of which are optimized through the use of a Proximal Policy Optimization (PPO)–based reinforcement learning model. The main goal of the proposed solution was to improve trading precision and dependability by minimizing false signals that are often produced by the use of a single indicator in the market conditions of high volatility and low liquidity.

The experimental results that have come from extensive backtesting on large-cap stocks prove that the SMAC-only strategy, although it is good in terms of detecting the trend direction, suffers from frequent false signals and the problem of excessive trading. The addition of the ATV as a volume confirmation tool has a remarkable effect on signal quality as it ensures that the price changes are accompanied by strong market participation. In addition, the incorporation of PCR reflects market sentiment based on derivatives, thereby synchronizing the trade decisions with the overall investors' expectations and consequently eliminating the errors that are caused by the sentiment.

Moreover, the PPO-based reinforcement learning has further strengthened the support of the framework through its provision of adaptive trade execution. The model learns optimal buy, sell, and hold actions from historical market behavior and thus it supports the improvement of timing for both entry and exit while keeping the risk-adjusted performance favorable. The proposed SMAC–ATV–PCR–PPO framework has been able to attain consistently higher profitability, better trade precision, and Sharpe and Sortino ratios that are superior to the ones obtained through the use of short-term, medium-term, and long-term trading horizons.

To sum up, the results verify that the integration of trend, volume, and sentiment indicators in a reinforcement learning approach leads to a strong and understandable automated trading system. The suggested approach has already reached a stage of maturity where it can be utilized in actual equity market trading and, moreover, it can broaden its application to other asset types like indices, derivatives, and commodities. Next, it might be a good idea to do some research on the addition of different sentiment sources, transaction cost modeling, and multi-asset portfolio optimization, which would contribute to further performance enhancement of the system.

Disclosure of Interests:
Conflict of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Informed Consent: This study did not involve human participants, and therefore, informed consent was not required.

Data source:
Source: Yahoo Finance, with feature engineering calculated on a Python platform
Link: https://finance.yahoo.com/quote/RELIANCE.NS/history/

References

1. Jaisson, T. (2022). Deep differentiable reinforcement learning and optimal trading. Quantitative Finance, 22(8), 1429–1443. https://doi.org/10.1080/14697688.2022.2062431
2. Kadia, A., Adhikary, S., Dey, R., Kar, A. (2025). “Deep Learning Based Stock Trading Strategies Using Leading Multi-Indicator Confirmations”. International Journal on Science and Technology (IJSAT). 16(3) IJSAT25037682 1-15. https://doi.org/10.71097/IJSAT.v16.i3.7682
3. Zhang, Y., Zohren, S., & Roberts, S. (2020). Deep reinforcement learning for trading. The Journal of Financial Data Science, 2(2), 25–40. https://doi.org/10.3905/jfds.2020.1.022
4. Adhikary S., Kadia A.,(2025), “Algorithmic Trading with a Combination of Advanced Technical Indicators – An Automation”, International Journal on Science and Technology (IJSAT). 16(3) IJSAT25037777 1-20, 2025, https://doi.org/10.71097/IJSAT.v16.i3.7777
5. Kadia, A., Dey, R., Kar, A. (2025). “Smart Stock Trading using an Advanced Combination of Technical Indicators with Volume Confirmation Integrated in Reinforcement Learning”. International Journal on Science and Technology (IJSAT). 16(3) IJSAT25037453 1-20. https://doi.org/10.71097/IJSAT.v16.i3.7453

6. Dey, R., Kassim, S., Maurya, S., Mahajan, R. A., Kadia, A., & Singh, M. (2024). A systematic literature review on the Islamic capital market: Insights using the PRISMA approach. Journal of Electrical Systems, 20(2s). https://doi.org/10.52783/jes.1571
7. Xiong, Z., Liu, D., Zhong, S., & Wu, C. (2023). Integrating technical indicators with deep Q-learning for automated stock trading. Applied Soft Computing, 131, 109852. https://doi.org/10.1016/j.asoc.2022.109852
8. Wang, Y., Zhang, Y., & Zheng, X. (2022). A hybrid reinforcement learning framework for financial signal trading. Quantitative Finance, 22(7), 1225–1241. https://doi.org/10.1080/14697688.2022.2037780
9. Bel Hadj Ayed, A., Loeper, G., & Abergel, F. (2018). Challenging the robustness of optimal portfolio investment with moving average-based strategies. Quantitative Finance, 19(1), 123–135. https://doi.org/10.1080/14697688.2018.1468080
10. Karaila, J., Baltakys, K., Hansen, H., Goel, A., & Kanniainen, J. (2024). Network analysis of aggregated money flows in stock markets. Quantitative Finance, 24(10), 1423–1443. https://doi.org/10.1080/14697688.2024.2409272
11. Zhou, J., Wang, J., & Li, F. (2021). Adaptive moving average crossover strategies with neural networks. Journal of Financial Markets, 53, 100588. https://doi.org/10.1016/j.finmar.2021.100588
12. Huang, J., Li, H., & Zhang, Y. (2023). Deep reinforcement learning for financial portfolio management: A survey. Expert Systems with Applications, 222, 119842. https://doi.org/10.1016/j.eswa.2023.119842
13. Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2020). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Systems with Applications, 42(1), 259–268. https://doi.org/10.1016/j.eswa.2019.07.025
14. Maqsood, M., Mehmood, T., & Farooq, A. (2022). Hybrid strategy using technical indicators and sentiment analysis for stock price prediction. Financial Innovation, 8(1), 43. https://doi.org/10.1186/s40854-022-00335-5
15. Li, S., Zhou, Z., & Wang, Y. (2021). Intelligent trading systems based on deep reinforcement learning: A systematic survey. IEEE Transactions on Neural Networks and Learning Systems, 32(11), 4798–4819. https://doi.org/10.1109/TNNLS.2020.3032395
16. Yu, Y., Huang, H., & Qin, Z. (2022). Reinforcement learning in financial market applications: Recent advancements and challenges. ACM Transactions on Intelligent Systems and Technology, 13(2), 1–27. https://doi.org/10.1145/3488510
17. Xiong, Z., Liu, D., Zhong, S., & Wu, C. (2023). Integrating technical indicators with deep Q-learning for automated stock trading. Applied Soft Computing, 131, 109852. https://doi.org/10.1016/j.asoc.2022.109852
18. Huang, J., Li, H., & Zhang, Y. (2023). Deep reinforcement learning for financial portfolio management: A survey. Expert Systems with Applications, 222, 119842. https://doi.org/10.1016/j.eswa.2023.119842
19. Jha, A., Maheshwari, S., Dutta, P., & Dubey, U. (2025). Optimizing financial modeling with machine learning: integrating particle swarm optimization for enhanced predictive analytics. Journal of Business Analytics, 8(3), 196–215. https://doi.org/10.1080/2573234X.2025.2470191

20. Batten, J. A., Szilagyi, P. G., & Wong, M. C. S. (2014). Stock Market Spread Trading: Argentina and Brazil Stock Indexes. Emerging Markets Finance and Trade, 50(sup3), 61–76. https://doi.org/10.2753/REE1540-496X5003S304
21. Kamble, A., & Patil, A. (2023). Volume-weighted reinforcement learning strategy for intraday stock trading. Procedia Computer Science, 218, 328–335. https://doi.org/10.1016/j.procs.2023.01.039
22. Wu, Y., & He, Q. (2023). Enhanced Technical Indicator Fusion with Reinforcement Learning in Stock Trading. Expert Systems with Applications, 213, 119002. https://doi.org/10.1016/j.eswa.2022.119002