Machine Learning Approach to Stock Market Prediction

Jing-Hao Ku1, a), Choo-Yee Ting1, 2, b), Goh Hui-Ngo1, 3, c) and Albert Quek1, 4, d)

1Faculty of Computing and Informatics, Multimedia University, Persiaran Multimedia, 63100, Cyberjaya, Malaysia.

2Centre for Big Data and Blockchain Technologies, CoE for Advanced Cloud, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Malaysia.

3Centre for Natural Language Processing, CoE for Artificial Intelligence, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Malaysia.

4Centre for Image and Vision Computing, CoE for Artificial Intelligence, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Malaysia.

*b) Corresponding author:* [*cyting@mmu.edu.my*](mailto:cyting@mmu.edu.my)

*a) 1201203310@student.mmu.edu.my*

*c) hngoh@mmu.edu.my*

*d) quek.albert@mmu.edu.my*

**Abstract.** Stock price prediction is crucial in the financial industry, helping investors make informed decisions, optimize portfolios, and improve profitability. Accurate forecasts are essential for risk management, capital allocation, and market efficiency. Conventional methods like technical analysis provide insight, but often fail to capture the complex, nonlinear patterns, and volatility of financial markets. Recently, machine learning has emerged as a powerful tool for more accurate and adaptive stock price predictions. This research proposes a novel machine learning framework that enhances conventional technical indicators by integrating predictive models as supplementary filters for Zero-Lag MACD trading signals. The framework was evaluated through comprehensive backtesting across different market capitalizations, including TSLA, GTLB, and HRMY, using hourly data over a five-month period. Key findings demonstrate that machine learning-enhanced strategies significantly outperformed conventional approaches across all tested securities. XGBoost performed the best overall across all stocks, with an average ROI of 40.75%, followed by LSTM with an average ROI of 37.09%. LSTM excelled with large-cap stocks, achieving a 90.12% ROI for TSLA. Linear Regression performed best for mid-cap stocks, delivering a 59.25% ROI for GTLB, while XGBoost proved most effective for small-cap stocks, achieving a 31.34% ROI for HRMY. The framework also achieved a substantial reduction in false signals, with machine learning models generating significantly fewer trading signals while producing higher returns. This research contributes a practical hybrid approach that maintains the interpretability of conventional technical indicators while enhancing their predictive accuracy through machine learning, demonstrating that even relatively simple models can substantially improve trading outcomes when properly integrated with established technical analysis frameworks.

# iNTRODUCTION

The stock markets play a key role in the economy and the world economy because they exchange equity and show how healthy the economy is [1]. Precise pricing projection is essential in investment plans and economic steadiness. [2, 3]. Stock market prediction refers to the application of multiple methods in estimating the future values of financial assets [4] which are impactful to the short-term and the long-term investment patterns [5, 6]. Stock market prediction is quite important because it minimizes risks and maximizes investment options in unstable financial atmospheres [4]. As an investor, regardless of the type of investment, it is necessary to know the movements of the market to remain profitable [6]. Nonetheless, this activity is always associated with certain challenges because of the volatility of the financial markets, which are also under the influence of diverse factors such as the state of the economy, the moods of traders and investors, reporting, and geopolitical factors [7, 8]. Errorful predictions may lead to losses of individual investors and influence the dynamics of the market to a greater extent [9, 10]. Stock market behavior is quite sophisticated and requires high-level prediction means. Machine learning has reflected immensely in this field where most traders have embraced the methods of machine learning in the financial sector [10, 11, 12, 13]. The latest developments are quite promising and addressing the issue of stock market prediction using algorithms that can be fed historical information and that can be adjusted to the current market changes [14, 15, 16]. Additionally, there are several key challenges in this area. A key challenge is that technical indicators are not alone enough to make proper stock market prediction, being rigid and not good in measuring the non-linear relationships existing in the stock market [17]. Those indicators rely too much on the lagging data and predefined strategies, which cannot take into consideration underlying non-linear correlations defining market dynamics. The given restrictions are further exacerbated by the fact that conventional trading strategies are mostly built on technical indicators and crude heuristics [12, 18]. Such techniques are prone to mistargeting and little flexibilities to changing market dynamics and thus hardly aid in real trading situations [5, 13, 15, 19]. Moreover, most advanced machine learning techniques also exhibit major drawbacks of interpretability as they are termed as black boxes where predictions are made without any comprehensible evidence of how the predictions were derived which is not suitable in a trading environment where traders must rely on the rationale behind investing advice [12]. When these limitations are put together, it is necessary to develop enhanced predictive techniques that will consider accuracy and transparency in financial projections.

The objectives of this study are:

1. To propose a machine learning approach to price and trend prediction.
2. To compare the performance of various machine learning models through backtesting.
3. To benchmark machine learning approach against conventional technical indicators through backtesting of trading strategies.

This research addresses critical limitations in current stock market prediction approaches, where conventional technical indicators often generate false signals, and advanced machine learning models lack interpretability for practical trading applications. Additionally, the comparative analysis of multiple machine learning models across different market capitalizations aims to provide practical insights for model selection based on stock characteristics. The study’s significance lies in bridging the gap between traditional technical analysis and modern predictive modeling, potentially offering a solution that combines familiar analytical frameworks with enhanced accuracy.

The practical significance extends to multiple stakeholder groups in the financial industry. Individual traders and institutional investors face ongoing challenges in developing reliable trading strategies that balance performance with transparency. This research could provide valuable insights for improving decision-making processes while maintaining regulatory compliance requirements that demand explainable trading logic. Furthermore, this study contributes to the academic understanding of hybrid approaches in financial forecasting, where the integration of conventional and machine learning methods remains underexplored. The findings could inform future research directions in computational finance and provide a foundation for developing more sophisticated trading systems that address both accuracy and interpretability concerns in financial markets.

# RELATED WORK

## Statistical Techniques in Stock Price Prediction

Statistical methods remain foundational in stock market prediction despite the emergence of more complex approaches. Recent studies show interesting patterns in research methods. Researchers use similar techniques across different projects. This creates a clear picture of current practices in the field. We examined multiple studies to understand these patterns. The results were quite revealing. Most researchers rely on established statistical methods for their work. Table 1 shows the main statistical approaches used in stock price prediction research. We focused on technical indicators that appear most often. These indicators help researchers analyze market trends effectively.

Simple Moving Average (SMA), as well as Exponential Moving Average (EMA) form Moving Averages (MA) trends because they smooth the fluctuations in prices [21]. These techniques help traders assess market directions[22]. Short-term moving averages help spot quick price changes. In contrast, long-term moving averages are usually better for following ongoing trends [13]. Another common tool is the Moving Average Convergence Divergence (MACD). It looks at how two moving averages relate to each other. This helps traders notice changes in momentum, which can signal shifts in direction [20]. It consists of the MACD Line, Signal Line, and Histogram, with crossovers providing valuable trading signals. That said, no indicator is perfect. Sometimes the signals can be unclear or even misleading. So, traders often use these tools with caution and combine them with other methods. Research has shown MACD-based strategies using GRU models demonstrated positive Annual Rate of Return across multiple stock exchanges and a Sharpe Ratio greater than 3 for 15 out of 18 stocks tested [20]. [15] achieved 75% accuracy using MACD individually, and 83.6% when combined with other technical indicators, while [16] incorporated MACD in their Enhanced Hybrid Trading System, demonstrating superior performance in returns and Sharpe ratio.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **TABLE 1.** Statistical techniques used by researchers | | | | | | | | | | |
| **Author** | **MA (SMA, EMA** | **MACD** | **RSI** | **Stochastic Oscillator** | **ATR** | **William's %R** | **CCI** | **Momentum** | **Bollinger Bands** | **OBV** |
| [2] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [9] |  | ✓ | ✓ | ✓ |  |  |  |  | ✓ |  |
| [10] | ✓ |  | ✓ | ✓ |  | ✓ | ✓ | ✓ |  |  |
| [13] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [15] | ✓ | ✓ | ✓ | ✓ |  | ✓ |  |  |  |  |
| [16] | ✓ | ✓ | ✓ |  |  |  |  |  |  |  |
| [17] | ✓ | ✓ | ✓ | ✓ |  | ✓ | ✓ | ✓ |  |  |
| [19] | ✓ | ✓ | ✓ | ✓ |  |  |  |  |  |  |
| [20] |  | ✓ |  |  |  |  |  |  |  |  |
| [21] | ✓ | ✓ | ✓ |  | ✓ |  | ✓ | ✓ |  | ✓ |
| [22] | ✓ | ✓ | ✓ |  |  |  |  |  |  | ✓ |
| [23] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

The Relative Strength Index (RSI) effectively identifies overbought or oversold conditions, with values above 70 or below 30 signaling potential reversals [21], with [23] achieving 85% accuracy in Pakistani markets and [14] successfully integrating it into ensemble models. The Stochastic Oscillator, comprising %K and %D lines, has proven valuable for detecting momentum shifts and market conditions, performing best in ranging markets [21] and improving prediction accuracy when combined with other indicators [13]. The Average True Range (ATR) measures market volatility by calculating the average of true ranges over a specified period [21], with [16] demonstrating improved performance metrics when integrated into trading systems. William’s %R has shown impressive results in predicting reversals, with accuracy rates of 78-85% across various models [23]. The Commodity Channel Index (CCI) identifies overbought or oversold conditions and detects potential trend reversals [21], with [13] achieving lower error rates through indicator combinations. Bollinger Bands consist of a moving average with upper and lower bands based on standard deviation [17], and have been incorporated into comprehensive predictive feature sets [10]. Volume-based indicators like On-Balance Volume (OBV) confirm price trends based on volume flow, with [22] successfully incorporating OBV into ensemble frameworks for predicting market indices, though noting potential reliability issues in volatile markets.

## Machine Learning in Stock Price Prediction

The evolution toward machine learning approaches in stock market prediction reflects the growing recognition of these statistical limitations. Machine learning has significantly advanced stock price prediction capabilities. Table 2 presents an overview of machine learning techniques commonly employed by researchers to address the challenges of stock price prediction.

[11] demonstrated LSTM’s superiority over Random Forest for Microsoft stock prediction through its ability to capture long-term dependencies, with dropout regularization effectively preventing overfitting. Extending this research, [24] showed Linear Regression substantially outperforming ARIMA across NSE-listed companies, with remarkably low error rates. [3] compared Linear Regression, Logistic Regression, and Random Forest across major tech companies, finding Linear Regression achieved R² values approaching 1, while Random Forest excelled in directional predictions. Market characteristics appeared to influence model effectiveness across different stocks, suggesting the importance of tailored approaches. Deep learning architectures have produced exceptional results, with [5] comparing CNN, RNN, LSTM, and GRU models while integrating multiple data sources. Their GRU model achieved 99.88% accuracy after five epochs, surpassing LSTM’s 99.39% with a lower MAE. [6] found SVM performed exceptionally well with lower Mean Absolute Percentage Errors compared to regression-based models, and LSTM when tested on major companies.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **TABLE 2.** Top common machine learning techniques used by researchers | | | | | | | | | | | | |
| **Author** | **Linear Regression** | **Logistic Regression** | **SVR** | **Random Forest** | **Gradient Boosting** | **XG Boost** | **Light GBM** | **Ada Boost** | **KNN** | **CNN** | **LSTM** | **GRU** |
| [3] | ✓ | ✓ |  | ✓ |  |  |  |  |  |  |  |  |
| [5] |  |  |  |  |  |  |  |  |  | ✓ | ✓ | ✓ |
| [6] | ✓ |  | ✓ |  |  |  |  |  |  | ✓ | ✓ |  |
| [9] | ✓ | ✓ | ✓ | ✓ |  | ✓ |  |  | ✓ |  | ✓ | ✓ |
| [11] |  |  |  | ✓ |  |  |  |  |  |  | ✓ |  |
| [12] |  |  |  | ✓ | ✓ |  |  |  |  |  |  |  |
| [14] |  |  |  | ✓ | ✓ | ✓ |  | ✓ |  |  |  |  |
| [18] | ✓ |  |  | ✓ | ✓ |  |  |  | ✓ |  |  |  |
| [21] |  | ✓ | ✓ | ✓ | ✓ |  |  |  | ✓ |  | ✓ |  |
| [24] | ✓ |  |  |  |  |  |  |  |  |  | ✓ |  |
| [25] |  |  | ✓ |  |  |  |  |  |  |  | ✓ |  |
| [26] | ✓ |  |  | ✓ |  | ✓ |  |  |  |  |  |  |
| [27] |  |  |  |  |  |  |  |  |  | ✓ | ✓ |  |
| [28] |  |  | ✓ | ✓ |  |  |  |  | ✓ |  | ✓ | ✓ |

[21] combined conventional candlestick charting with modern ensemble techniques across multiple models, finding Random Forest and Gradient Boosting achieved F1 scores exceeding 0.80 for certain patterns in the Chinese market. [12] evaluated ensemble methods using the Quantopian trading simulator, with Gradient Boosting achieving 57% accuracy while maintaining lower volatility. [14] found XGBoost most effective for predicting Indian bank stock returns, outperforming Random Forest and AdaBoost. [25] showed LSTM consistently achieved lower MAPE values than SVR, particularly for the S&P 500 index at 0.75%. Contrary to expectations of complex model superiority, [18] found Linear Regression achieving perfect R² values and lowest RMSE across all tested stocks compared to KNN, Random Forest, and Gradient Boosting. For individual stocks, [26] demonstrated XGBoost Regressor’s superiority with 98.65% accuracy, outperforming both Linear Regression and Random Forest. [27] found that LSTM and MLP models were particularly effective at enhancing the predictive power of conventional technical indicators across major markets, demonstrating improved trading signal success rates and Sharpe ratios. [9] evaluated fifteen different models, finding ensembles combining Random Forest, XGBoost, and LSTM produced the most accurate predictions for Indian stocks. Finally, [28] compared machine learning approaches with an ensemble method for Reliance stock, with LSTM achieving the highest R² value of 0.838 compared to the hybrid approach, demonstrating that model selection should align with specific stock characteristics.

Based on the literature review, several key gaps persist in stock market prediction research. Despite progress in machine learning, its application in trading strategies is still limited [14, 27]. Moreover, many studies lack realistic backtesting frameworks that reflect actual trading conditions, reducing the real-world applicability of their findings. The opaque, "black box" nature of advanced ML models also raises interpretability issues, making them less trustworthy for practitioners [4]. These problems raise the importance of methods with better practical use, and model transparency.

# METHOD

This section describes the approaches used to collect, preprocess financial data, construct trading strategies, optimize model parameters, and evaluate system performance through comprehensive metrics and backtesting.

## Data Source

The framework uses yfinance to collect historical stock market data from Yahoo Finance for selected stocks, including TSLA (Tesla, Inc.), GTLB (GitLab Inc.), and HRMY (Harmony Biosciences Holdings, Inc.). It retrieves key metrics such as opening prices, closing prices, highs/lows, volumes, and adjusted closing prices. The framework supports multiple time series with different intervals, such as daily and hourly, enabling analysis across various trading timeframes.

## Data Preprocessing

Data preprocessing transforms raw financial data into a format suitable for machine learning models through several critical steps. The framework implements time index handling that maps different time intervals to appropriate pandas frequency strings (15T, 30T, H, B) to meet the requirements of the Darts time series analysis library. Missing data points are addressed through forward filling, ensuring temporal continuity without artificial gaps. MinMaxScaler is applied separately to target and covariate features following a careful sequence to prevent data leakage: first splitting data into 80-20 training-testing sets without random shuffling, then computing normalization parameters solely from training data, and finally applying these parameters to both datasets. The feature engineering phase selects close prices as the target variable and utilizes open, high, low, and volume as covariates, deliberately excluding adjusted close price to prevent data leakage and overfitting.

## Trading Strategy Construction

The construction of the trading strategy will incorporate and compare five machine learning models: Linear Regression, Random Forest, XGBoost, LSTM, and GRU, and will be evaluated through historical backtesting. The backtesting capital will be set at $10,000, with the strategy buying as many shares as possible based on available capital when a buy signal is generated, and selling all the shares held when a sell signal occurs. The backtesting assumes a transaction fee of 1% for each trade, impacting the overall performance calculation. Going on this, there are five different trading strategies of the Zero-Lag MACD founded on the association with the signal line and zero axis. These are Buy Above Sell Above (BASA), Buy Below Sell Above (BBSA), Buy Above Sell Below (BASB), Buy Below Sell Below (BBSB) and Histogram Trend Reversal (HTR), each uses different conditional logic in order to generate a signal. Moreover, the system has an optimized parameter strategies that loop a combination of fast, slow, and signal parameters in an effort to perfect each strategy. The system allows varying different sets of parameters to make sure that the most productive combinations of parameters are used so that ROI is maximized. To minimize false signals, the machine learning model is implemented into the additional signal filter so that the filtering method is to validate signals based on Zero-Lag MACD with one-step-ahead price forecasts. The signal of buy is verified when the forecasted priced at t+1 is at higher value than the current price at t and sells signal is verified when the value at the forecasted price is lower than the current price to improve the conventional Zero-Lag MACD signal by adding a predictive aspect to it without compromising feasibility of simulation.

## Hyperparameter Tuning

Hyperparameter tuning phase adopts a grid strategy to search the predetermined parameter spaces systematically across all models of machine learning. The described process involves employing the output of Grid Search, which aims at entirely considering every set of hyperparameter combinations. The term cross validation refers to methods of combining cross validation directly into the tuning process by measuring how the model performs on other subsets of the data to guarantee robust generalization and avoid over-fitting of the training data by making sure that the models will spot real patterns and not just noise in the training data.

## Evaluation

The evaluation phase refers to a sophisticated assessment system based on which a performance of all the five machine learning models is analyzed through various perspectives. The backtesting engine is an engine that tests trading performance of the selected strategy with management about trading fees, position sizing, and keeping track of the trade log. It also computes some aspects of performance such as ROI, maximum profit and loss per trade and total numbers of trades.

# FINDINGS

This paper presents the results and analysis derived from the implementation of the proposed methodologies in this study. The backtesting of the proposed machine learning-improved trading strategies was performed in a backtesting process comprising all the three securities of varying market cap. The backtesting was carried out between October 1, 2024 to March 1, 2025 and time frame was hourly to cover different market conditions such as bull and bear market. This analysis allows strong testing of the suggested methods in a variety of market situations and industry-related dynamics. Table 3 shows a detailed comparison of conventional Zero-Lag MACD and the five machine learning enhancements of the algorithm on all the securities and strategies tested. These findings depict that the combination of machine learning models as additional signal filters showed effectiveness through enhancing trading performance.

Table 3 indicates that all the five machine learning-enriched trading strategies provide superior performance over the traditional benchmark methods using various market cap and sector. The use of machine learning models as trading signals enhancement mechanism shows significant changes in returns of investment in all securities tested.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **TABLE 3.** Performance comparison of conventional and machine learning-enhanced trading strategies | | | | | | | | | | | | | | | | | | | | | | | | | |
| **Stock Name** | **Strategy** | **Zero-Lag MACD** | | | | **Linear Regression Enhanced Zero-Lag MACD** | | | | **Random Forest Enhanced Zero-Lag MACD** | | | | **XG Boost Enhanced Zero-Lag MACD** | | | | **LSTM Enhanced Zero-Lag MACD** | | | | **GRU Enhanced Zero-Lag MACD** | | | |
| **Frequency** | **Max Profit ($)** | **Max Loss ($)** | **ROI(%)** | **Frequency** | **Max Profit ($)** | **Max Loss ($)** | **ROI(%)** | **Frequency** | **Max Profit ($)** | **Max Loss ($)** | **ROI(%)** | **Frequency** | **Max Profit ($)** | **Max Loss ($)** | **ROI(%)** | **Frequency** | **Max Profit ($)** | **Max Loss ($)** | **ROI(%)** | **Frequency** | **Max Profit ($)** | **Max Loss ($)** | **ROI(%)** |
| TSLA | BASA | 18 | 241.77 | -209.73 | 14.99 | 2 | 257.22 | -79.86 | 69.97 | 3 | 238.86 | -30.91 | 85.67 | 1 | 152.62 | 152.62 | 64.53 | 4 | 233.63 | 15.18 | 132.84 | 3 | 165.34 | -121.91 | 56.55 |
| TSLA | BBSA | 13 | 255.69 | -217.47 | 20.13 | 3 | 157.14 | -22.10 | 67.87 | 5 | 196.92 | -53.86 | 83.17 | 3 | 157.14 | 2.07 | 80.33 | 3 | 255.46 | -14.17 | 111.83 | 5 | 165.38 | -87.19 | 88.28 |
| TSLA | BASB | 2 | 197.19 | 18.85 | 90.87 | 1 | 237.60 | 237.60 | 103.89 | 1 | 237.66 | 237.66 | 103.92 | 2 | 201.85 | 17.74 | 92.40 | 2 | 202.62 | 14.02 | 88.92 | 3 | 237.66 | -52.33 | 94.89 |
| TSLA | BBSB | 5 | 212.56 | -157.26 | 53.81 | 1 | 211.20 | 211.20 | 81.68 | 2 | 237.82 | 0.48 | 100.07 | 4 | 235.89 | -35.04 | 96.56 | 3 | 238.69 | -72.66 | 67.02 | 4 | 237.82 | -37.03 | 87.17 |
| TSLA | HTR | 18 | 244.29 | -212.50 | -8.85 | 11 | 185.88 | -171.27 | 11.81 | 7 | 191.80 | -52.08 | 51.09 | 9 | 169.72 | -149.30 | 50.25 | 8 | 214.81 | -187.43 | 49.99 | 11 | 237.78 | -158.17 | 66.06 |
| GTLB | BASA | 6 | 19.87 | -9.95 | 15.14 | 2 | 21.79 | 11.22 | 65.46 | 3 | 20.69 | -6.28 | 40.73 | 3 | 20.69 | 0.59 | 58.79 | 7 | 19.91 | -16.07 | 31.97 | 6 | 23.24 | -10.64 | 39.40 |
| GTLB | BBSA | 12 | 20.17 | -22.37 | 12.52 | 4 | 24.04 | -9.42 | 62.81 | 4 | 14.54 | -7.62 | 41.23 | 5 | 19.68 | -4.17 | 59.62 | 5 | 19.73 | -3.32 | 60.50 | 12 | 20.35 | -13.85 | 44.40 |
| GTLB | BASB | 2 | 18.05 | 2.72 | 34.84 | 2 | 18.26 | 11.22 | 57.31 | 2 | 20.61 | 5.33 | 48.22 | 4 | 16.73 | -3.09 | 47.63 | 2 | 22.28 | 4.73 | 50.37 | 4 | 20.75 | -8.70 | 44.57 |
| GTLB | BBSB | 8 | 20.67 | -14.06 | 21.17 | 2 | 21.39 | -0.10 | 67.49 | 3 | 20.64 | -9.41 | 42.24 | 5 | 21.70 | -3.25 | 73.47 | 8 | 21.56 | -9.17 | 63.56 | 8 | 18.81 | -8.06 | 65.92 |
| GTLB | HTR | 20 | 23.82 | -22.10 | -12.54 | 9 | 23.68 | -14.22 | 44.19 | 13 | 24.50 | -16.70 | 41.23 | 7 | 18.29 | -4.49 | 36.92 | 14 | 23.88 | -20.98 | 23.03 | 12 | 20.25 | -12.95 | 32.91 |
| HRMY | BASA | 4 | 5.12 | -3.33 | -1.48 | 3 | 7.12 | -2.19 | 25.41 | 2 | 6.60 | -0.14 | 14.79 | 3 | 8.36 | 1.08 | 45.74 | 7 | 7.56 | -4.71 | 15.16 | 9 | 8.36 | -5.41 | 15.38 |
| HRMY | BBSA | 10 | 5.90 | -5.86 | 0.77 | 5 | 7.23 | -5.11 | 42.71 | 2 | 6.46 | -2.92 | 17.90 | 3 | 8.81 | 0.49 | 36.02 | 7 | 7.79 | -4.1 | 33.42 | 7 | 7.79 | -4.18 | 33.42 |
| HRMY | BASB | 1 | 1.35 | 1.35 | 2.02 | 1 | 5.60 | 5.60 | 14.14 | 2 | 5.73 | -0.38 | 14.99 | 4 | 8.17 | -2.91 | 31.48 | 8 | 8.47 | -4.67 | 13.85 | 7 | 7.31 | -4.08 | 11.25 |
| HRMY | BBSB | 5 | 5.48 | -.366 | 0.25 | 1 | 4.42 | 4.42 | 10.31 | 1 | 5.65 | 5.65 | 14.30 | 1 | 4.40 | 4.40 | 10.70 | 4 | 7.18 | -3.77 | 22.26 | 4 | 7.18 | -3.77 | 22.26 |
| HRMY | HTR | 17 | 7.89 | -7.36 | -37.64 | 6 | 8.10 | -5.54 | 34.07 | 8 | 6.86 | -6.11 | -5.32 | 5 | 8.43 | -4.00 | 32.78 | 21 | 7.54 | -5.55 | -2.40 | 15 | 7.63 | -5.97 | -7.72 |

The model brought improvement since the strategies had better returns with various market capitalization, not only with large-cap technology stocks TSLA but also mid-cap stocks GTLB and small-cap healthcare investments HRMY. This has been found in all the various sectors and shows that models would fit dynamically with the changes in the market and industry specific aspects. The probable success of machine learning methods compared to the traditional methods is clear when the averaging of the results of the various trading strategies of 5 stocks is performed. In the case of TSLA, LSTM proved to be the best among them by achieving an average ROI of 90.12% which was a stunning 55.93% better than the ROI of Zero-Lag MACD of 34.19%. It is interesting to note that LSTM was able to perform better in addition to emitting 64.3% fewer signals, which has enhanced signal accuracy of LSTM to a great extent and achieved remarkable noise and false positive cut down as well as compared to the traditional Zero-Lag MACD. Random forest came next as the second-best performer on TSLA to achieve an 84.78% ROI status, exhibiting the strength of tree-based ensemble algorithms when dealing with large caps and tech stocks. In mid-cap GTLB, it can be said that Linear Regression was best with the average ROI of 59.25%, which is 45.02% more than Zero-Lag MACD with a 14.23% ROI and fueled the accuracy of signal frequency generated by 60.4%. XGBoost was right behind the heels with 55.29% ROI, which implies that both non-linear and linear methods are effective in mid-cap stocks, provided they are used optimally. In the case of small-cap HRMY, XGBoost has proven to be very effective with 31.34% ROI compared to the -7.22% of Zero-Lag MACD, which performs dramatically better with a decrease in signal frequency caused accuracy by 56.8%. Linear regression was also a success in the case of HRMY with an ROI of 25.33%. which mean that even basic models can work out profitable trends in small-cap securities. In a nutshell, XGBoost and LSTM performed consistently in all the stocks with an average ROI of 40.75% and 37.09%, respectively. LSTM performance with large-cap stocks was also superior due to spectacular performance with TSLA, whereas, XGBoost and Linear Regression performed well with small-cap and mid-cap stocks lending to HRMY and GTLB. It indicates that albeit the machine learning tools can reach better overall performance, one should select the model depending on the market capitalization and nature of a stock.

It is this enhancement in the quality of signals that seeks to solve one of the inherent weakness of traditional technical indicators, their propensity to provide random or false positive signals in aggressive market environments. The improved signal accuracy resulted into better time in trade and also played a greater role in superior ROI results. These findings can be considered as evidence that the use of machine learning to generate trading signals gives a significant boost in dealing with complex and often nonlinear relationships that are inherent in contemporary financial markets when conventional trading methods alone are used. The confirmation of Zero-Lag MACD signals using prices forecasted one step ahead helped the system to integrate the interpretability of traditional technical indicators and previous tracks found in the literature review overcoming the so-called black box limitations in which the systems are said to have. This hybrid approach maintains the familiarity and transparency of conventional trading strategies while enhancing their performance through data-driven validation.

# CONCLUSION

This research addressed the challenge of improving stock market prediction through a machine learning approach. The study achieved its objectives by comparing multiple machine learning models to identify the best performing models for price prediction, developing optimized trading strategies that integrate these predictions with technical indicators, and establishing a comprehensive benchmarking framework that demonstrated clear performance improvements over conventional methods. The findings reveal that machine learning models consistently outperformed conventional approaches across all tested securities. LSTM excelled with large-cap stocks (TSLA), achieving a 90.12% ROI while generating 64.3% fewer signals than conventional methods. Linear Regression performed best for mid-cap stocks (GTLB), with a 59.25% ROI, and XGBoost proved most effective for small-cap stocks (HRMY), with a 31.34% ROI compared to Zero-Lag MACD's negative return. Overall, XGBoost and LSTM demonstrated the best performance across all stocks, with average ROIs of 40.75% and 37.09% respectively. Despite these promising results, several limitations warrant attention in future work. The present version is more inclined to only one machine learning model, so the chances of absorbing more advanced market patterns are reduced. The next stage of research will focus on the ensemble approach with the use of multiple models that predict performance in a better way. Besides, the existing trading strategies are effective but more complex strategies, which consider adaptive position sizing and risk management should be developed. Last, but not least the translation of a backtesting model into a live trading environment is an essential next step to indicate performance in the real world market under real-world execution constraints and latency considerations. Finally, this study provides a basis of improving the ability of predicting the stock market using machine learning and still staying close to the benefits of interpretability of the traditional methods. The future research in which these limitations could be extended further has a high chance of practical implementation in a real-life trading situation.

# References

1. Shah, A. Jain, R. Gupta, K. Bhowmick, and N. Patil, “International journal of intelligent systems and applications in engineering stock price prediction using ensemble model and sentiment analysis,” (2024).
2. Yang, Y. Wang, and X. Li, “Prediction of stock price direction using the lasso-LSTM model combines technical indicators and financial sentiment analysis,” PeerJ Computer Science **8**, e1148 (2022).
3. T. Shen, “Machine learning applications in stock price prediction,” Advances in Economics, Management and Political Sciences **94**, 257–267 (2024).
4. A. Das, K. Jain, D. Majumder, and R. Karmakar, “An ensemble learning paradigm for efficient stock market predictions,” in 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT) (IEEE, 2024).
5. A. Ismailova, Z. Beldeubayeva, K. Kadirkulov, Z. Doumcharieva, A. Konyrkhanova, D. Ussipbekova, A. Aripbayeva, and D. Yesmukhanova, “Forecasting stock market prices using deep learning methods,” International Journal of Electrical and Computer Engineering **14**, 5601–5611 (2024).
6. Q. Chen, “Stock market prediction using machine learning,” (2023), pp. 458–465.
7. I.K. Nti, A.F. Adekoya, and B.A. Weyori, “A systematic review of fundamental and technical analysis of stock market predictions,” Artificial Intelligence Review **53**, 3007–3057 (2020).
8. P.H. Vuong, T.T. Dat, T.K. Mai, P.H. Uyen, and P.T. Bao, “Stock-price forecasting based on XGBoost and LSTM,” Computer Systems Science and Engineering **40**, 237–246 (2022).
9. G. Sonkavde, D.S. Dharrao, A.M. Bongale, S.T. Deokate, D. Doreswamy, and S.K. Bhat, “Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications,” (2023).
10. M. Nabipour, P. Nayyeri, H. Jabani, S. Shahab, and A. Mosavi, “Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis,” IEEE Access **8**, 150199–150212 (2020).
11. B.K. Gupta, S. Tiwari, S. Bhatnagar, Shalu, Y. Singh, and T. Ranjan, “Stock price prediction using ML algorithms,” International Journal for Research in Applied Science and Engineering Technology **11**, 5426–5432 (2023).
12. R. Saifan, K. Sharif, M. Abu-Ghazaleh, and M. Abdel-Majeed, “Investigating algorithmic stock market trading using ensemble machine learning methods,” Informatica (Slovenia) **44**, 311–325 (2020).
13. G. Alfonso and D.R. Ramirez, “A nonlinear technical indicator selection approach for stock markets: Application to the Chinese stock market,” Mathematics **8**, 1301 (2020).
14. S. Mohapatra, R. Mukherjee, A. Roy, A. Sengupta, and A. Puniyani, “Can ensemble machine learning methods predict stock returns for Indian banks using technical indicators?” Journal of Risk and Financial Management **15**, 350 (2022).
15. M.C. Lee, J.W. Chang, J.C. Hung, and B.L. Chen, “Exploring the effectiveness of deep neural networks with technical analysis applied to stock market prediction,” Computer Science and Information Systems **18**, 401–418 (2021).
16. A.F. Kamara, E. Chen, and Z. Pan, “An ensemble of a boosted hybrid of deep learning models and technical analysis for forecasting stock prices,” Information Sciences **594**, 1–19 (2022).
17. Y. Xu, C. Yang, S. Peng, and Y. Nojima, “A hybrid two-stage financial stock forecasting algorithm based on clustering and ensemble learning,” Applied Intelligence **50**, 3852–3867 (2020).
18. D.S.A. Elminaam, A.E. Tanany, M.A. Salam, and M.A.E. Fattah, “CPSMP\_ML: Closing price prediction of stock market using machine learning models,” in 2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC) (IEEE, 2022), pp. 251–255.
19. M. Agrawal, P.K. Shukla, R. Nair, A. Nayyar, and M. Masud, “Stock prediction based on technical indicators using deep learning model,” Computers, Materials and Continua **70**, 287–304 (2021).
20. A.S. Saud and S. Shakya, “Technical indicator empowered intelligent strategies to predict stock trading signals,” Journal of Open Innovation: Technology, Market, and Complexity **10**, 100398 (2024).
21. Y. Lin, S. Liu, H. Yang, and H. Wu, “Stock trend prediction using candlestick charting and ensemble machine learning techniques with a novelty feature engineering scheme,” IEEE Access **9**, 101433–101446 (2021).
22. I.K. Nti, A.F. Adekoya, and B.A. Weyori, “A comprehensive evaluation of ensemble learning for stock-market prediction,” Journal of Big Data **7**, 99 (2020).
23. H. Raza and Z. Akhtar, “Predicting stock prices in the Pakistan market using machine learning and technical indicators,” Modern Finance **2**, 46–63 (2024).
24. A.M. Kushwaha, R.R. Malhotra, N.N. Manchhani, and V.G. Narwani, “Stock price prediction using machine learning models: A study of NSE stocks,” International Journal for Research in Applied Science and Engineering Technology **12**, 1642–1653 (2024).
25. G. Bathla, “Stock price prediction using LSTM and SVR,” in PDGC 2020 – 6th International Conference on Parallel, Distributed and Grid Computing (IEEE, 2020), pp. 211–214.
26. P. Durga and T. Sudhakar, “Machine learning classification methods for stock market forecasting,” in International Conference on Innovative Data Communication Technologies and Application (ICIDCA 2023) (IEEE, 2023), pp. 14–21.
27. Y. Shi, B. Li, W. Long, and W. Dai, “Method for improving the performance of technical analysis indicators by neural network models,” Computational Economics **59**, 1027–1068 (2022).
28. M. Rehman, M. Fuzail, M.K. Abid, and N. Aslam, “Financial prices prediction of stock market using supervised machine learning models,” VFAST Transactions on Software Engineering **11**, 1–10 (2023).