Market Forecasting Using LSTM-ARIMA Model with MACD Decomposition

Teng-Chih Yu [1] and Jian-Jiun Ding [2*]
Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan
E-mail: r12942126@ntu.edu.tw [1] and jjding@ntu.edu.tw [2*], Tel: +886-2-33669652

Abstract—**Market forecasting plays a critical role in economic signal processing. In this study, a model that applies the moving average convergence divergence (MACD) indicator together with learning-based architectures is proposed for market forecasting. First, we propose a novel MACD decomposition method to preprocess the data. By employing the proposed decomposition approach, the high- and low-frequency components are effectively separated. These components correspond to the short-term and the long-term information in the market, respectively. Then, two well-known models in economic forecasting, LSTM and ARIMA models, are employed to perform next-day predictions using only past closing prices. The proposed algorithm well integrates the merits of learning-based techniques and classical regression model. Despite of relatively less complexity, the proposed approach achieves promising performance and demonstrates adaptability across a variety of economic signals.**

I. Introduction

In the stock market, technical indicators, such as the moving average convergence divergence (MACD) and Relative Strength Index (RSI), are commonly proposed by experts and economists based on empirical rules. These indicators are widely used to identify entry points and evaluate the strength of stock trends. In the domain of machine learning, long short-term memory (LSTM)-based models have become the most prevalent approach for stock price prediction. Leveraging this, we employ the MACD indicator [1] from economic analysis to decompose the raw time series into high- and low-frequency components. These components are then individually predicted using LSTM [2] and ARIMA [3] models. By applying the above methodology, even when using only a one-dimensional time series (the closing price), enabling a more granular analysis of stock price movements.

In this paper, we propose an approach that integrates tools from the field of economics with machine learning techniques. The key contribution of our work lies in utilizing the MACD indicator to facilitate more efficient data training.

Long-term (low-frequency) features: For the low-frequency component, we utilize the Long Short-Term Memory (LSTM) model, a type of recurrent neural network (RNN) [4] specifically designed to capture long-term dependencies in sequential data. LSTMs leverage memory cells and gating mechanisms (forget, input, and output gates) to selectively store, update, and utilize information, making them particularly effective for time series forecasting tasks. Due to its widespread application across numerous tasks, the potential of LSTM remains a topic of active discussion in recent studies [5].

Short-term (high-frequency) features: Due to the strong performance of the ARIMA model in short-term prediction, it is well-suited for handling the high-frequency component. The ARIMA model consists of three components: autoregressive (AR), integrated (I), and moving average (MA). A key requirement for ARIMA is that the input data must be stationary, meaning its mean and variance remain constant over time. The integrated component of ARIMA addresses non-stationary means, while a logarithmic transformation is often applied to ensure variance stationarity.

The ARIMA model is parameterized by three values p, d, and q, where p represents the number of lag observations included in the autoregressive term, q indicates the number of lagged forecast errors in the moving average term, and d denotes the order of differencing applied to achieve stationarity. In the signal with trend, typically set to $d = 1$.

To optimize the ARIMA model parameters, we employ the Akaike information criterion (AIC) [6]. AIC balances model complexity and goodness of fit, with lower AIC values indicating better model performance. The ARIMA model is an active subject of research and discussion in the context of short-term forecasting [7].

This paper systematically introduces our proposed method, detailing the application of the MACD indicator to preprocess the data and the design of our model architecture. Furthermore, we conduct a comparative analysis with several commonly used models to evaluate the effectiveness and performance of our approach.

The **primary contribution** of this paper is its novel high–low frequency decomposition framework, which preprocesses real-world market indicators through mathematical techniques to yield frequency components that implicitly reflect the behavior of traders and can be applied across diverse economic signals. Building on this decomposition, we tailor the choice of regression model to the specific characteristics of each frequency component, thereby enhancing predictive accuracy.

979-8-3315-7206-8/25/$31.00 ©2025 IEEE

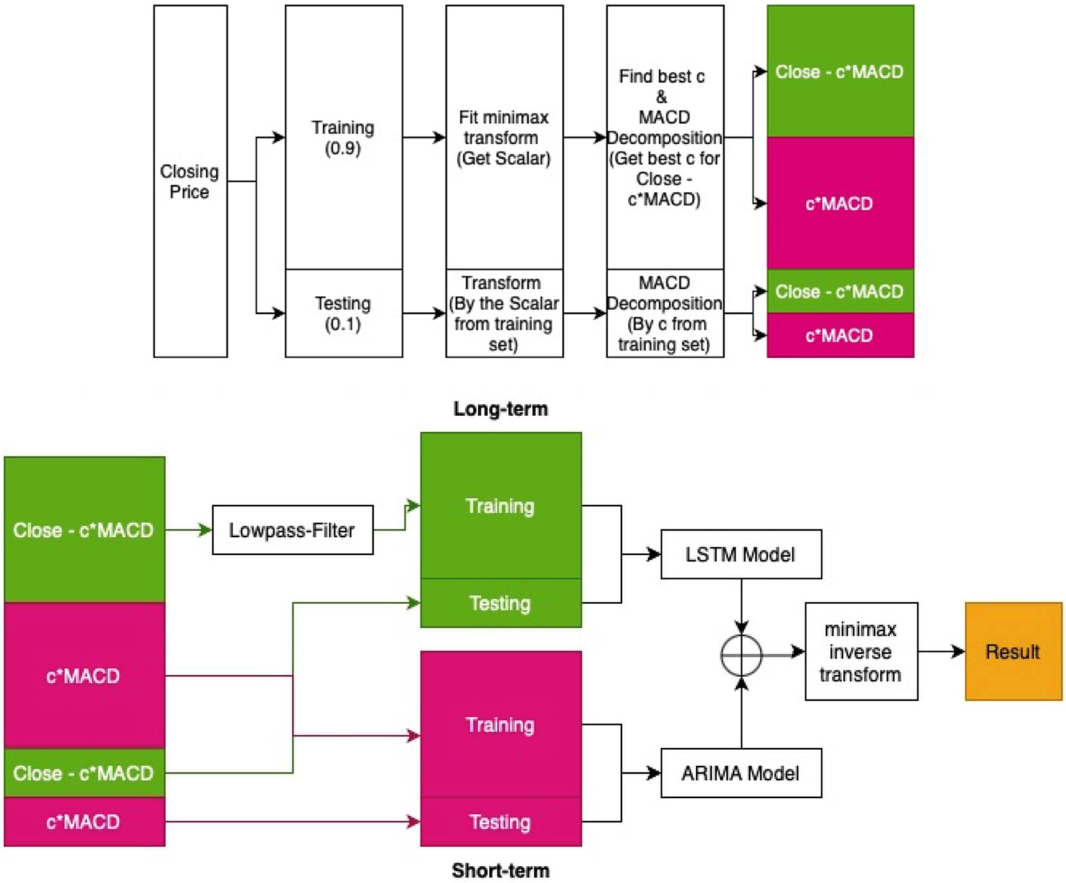

Fig. 1 Flow diagram of the proposed MACD-LSTM-ARIMA market prediction model.

Moreover, by combining advanced machine-learning methods with classical, interpretable regression models, we exploit the complementary strengths of both approaches to achieve robust and explanatory forecasting performance.

II. Proposed Method

The proposed method can be divided into two parts: (I) **Preprocessing** section and (II) **Model** section, based on the flow diagram, as in Fig. 1.

A. *Min-Max Normalization*

First, we apply a Min-Max scaling to the raw time-series data. This transformation serves to stabilize the input range during model training, in machine learning, it is common to use min max transform to stabilize training [8]. To prevent any look-ahead bias, we fit the Min-Max scaler exclusively on the training set's observed range. By performing this operation, we ensure that all subsequent scaling of validation and test data relies solely on information available at training time.

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

B. *MACD Indicator*

The MACD (Moving Average Convergence Divergence) is one of the most widely used economic indicators, introduced by Gerald Appel in the 1970s. It is defined as the difference between the short-term Exponential Moving Average (EMA) and the long-term EMA. The EMA at time t is defined as follows:

$$EMA_t = P_t \bullet K + EMA_{t-1} \bullet (1 - K), \quad (2)$$

where P_t is the price at time t, and for n-period EMA, n denotes the number of time steps considered. A larger n results in a smoother average by giving more weight to past observations, our first EMA will be generated until we obtain n data points, and it K can be defined as

$$K = 2/(n + 1). \quad (3)$$

After we obtain the EMA, we can calculate the MACD as following:

$$MACD_t = EMA_t^a - EMA_t^b, \quad (4)$$

where a and b represent the period of EMA, and $a < b$. In this paper, a and b are chosen as 12 and 26, which are the pair most used in the economic field. Up to the present day, the MACD technical indicator continues to be widely studied and utilized for trading purposes [9].

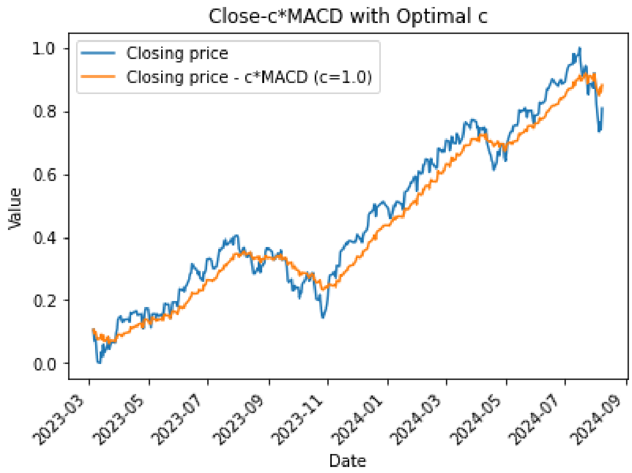

Fig. 2 The S&P500 closing price and its low frequency part decided by MACD-decomposition.

C. *MACD Decomposition with Logarithm*

A key feature of the MACD decomposition is its ability to preserve the critical high-frequency components present in market data. This arises because many traders rely on technical indicators to guide their buy and sell decisions, and those trading actions, in turn, exert a direct influence on market prices.

We first apply a logarithmic transformation to the entire 1-dimensional time series. This transformation reduces the scale of the original data and thereby mitigates the increased variance caused by large-magnitude values, and the variance can be stabilized by logarithmic transformation [10]; The value obtained here is the natural logarithm of the closing price $ln(Closing\ price)$. Furthermore, it also has its own MACD value, denoted here as $MACD^*$. As a result, the selected c more closely aligns with our objectives. It should be noted that the logarithmic transformation is used exclusively during the c-optimization phase and does not participate in the actual signal decomposition.

$$minimize_c\ Var(ln(Closing\ price) - c \bullet MACD^*), \qquad (5)$$

where c is a constant, $\hat{c}$ is its optimal value, and Var is the variance of $\ln(Closing\ price)$ minus c times the MACD of $ln(Closing\ price)$. The parameter c can be optimized by the partial derivative of the variance and solving a linear equation.

In this stage, the optimal coefficient $\hat{c}$ is applied to both the closing-price $Closing\ price$ and its corresponding MACD series $MACD$. The resulting decomposition produces a low-frequency component

$$L_t = \ Closing\ price - \hat{c} \bullet MACD \qquad (5)$$

and a high-frequency component

$$H_t = \ \hat{c} \bullet MACD. \qquad (6)$$

Fig. 2 illustrates the low-frequency component extracted from the decomposition of the closing-price series and its low

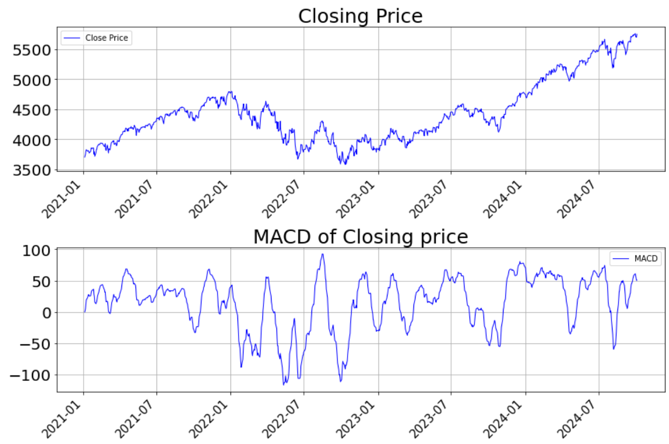

Fig. 3 The S&P500 closing price and its MACD value.

frequency component L_t. As shown, short-term fluctuations are substantially attenuated.

D. *Smoothing of the Low-Frequency Component*

In this stage, since the key high-frequency information has already been isolated, it is sensible to filter out any residual high-frequency noise from the low-frequency component, thereby enhancing stability during the training phase. Therefore, a low-pass filter is applied to the low-frequency component of the training set to enhance stability during the training phase.

E. *Model Selection for Long Term Signal: LSTM*

At this stage, we have successfully obtained both the low-frequency and high-frequency components. Because LSTM can retain information from distant past observations, it was chosen as the regression model for the low-frequency component. Moreover, by applying a low-pass filter to the low-frequency series in the preprocessing stage, the training process achieves greater stability, Furthermore, a complex model architecture is unnecessary.

F. *Model Selection for Short term Signal: ARIMA*

The ARIMA model is selected for the high-frequency component because it conditions its forecasts solely on the preceding pp days of observations, making it ideally suited for short-term prediction. Moreover, inspection of the MACD waveform Fig. shows that it oscillates symmetrically around zero and, since MACD amplitude is inherently linked to trend magnitude, its variance remains approximately constant over time. These properties exactly satisfy ARIMA's stationarity requirements—constant mean and variance—so that the high-frequency signal H_t, which is derived from MACD, unlike a trend-driven series, does not require differencing to meet the stationarity assumptions of the model.

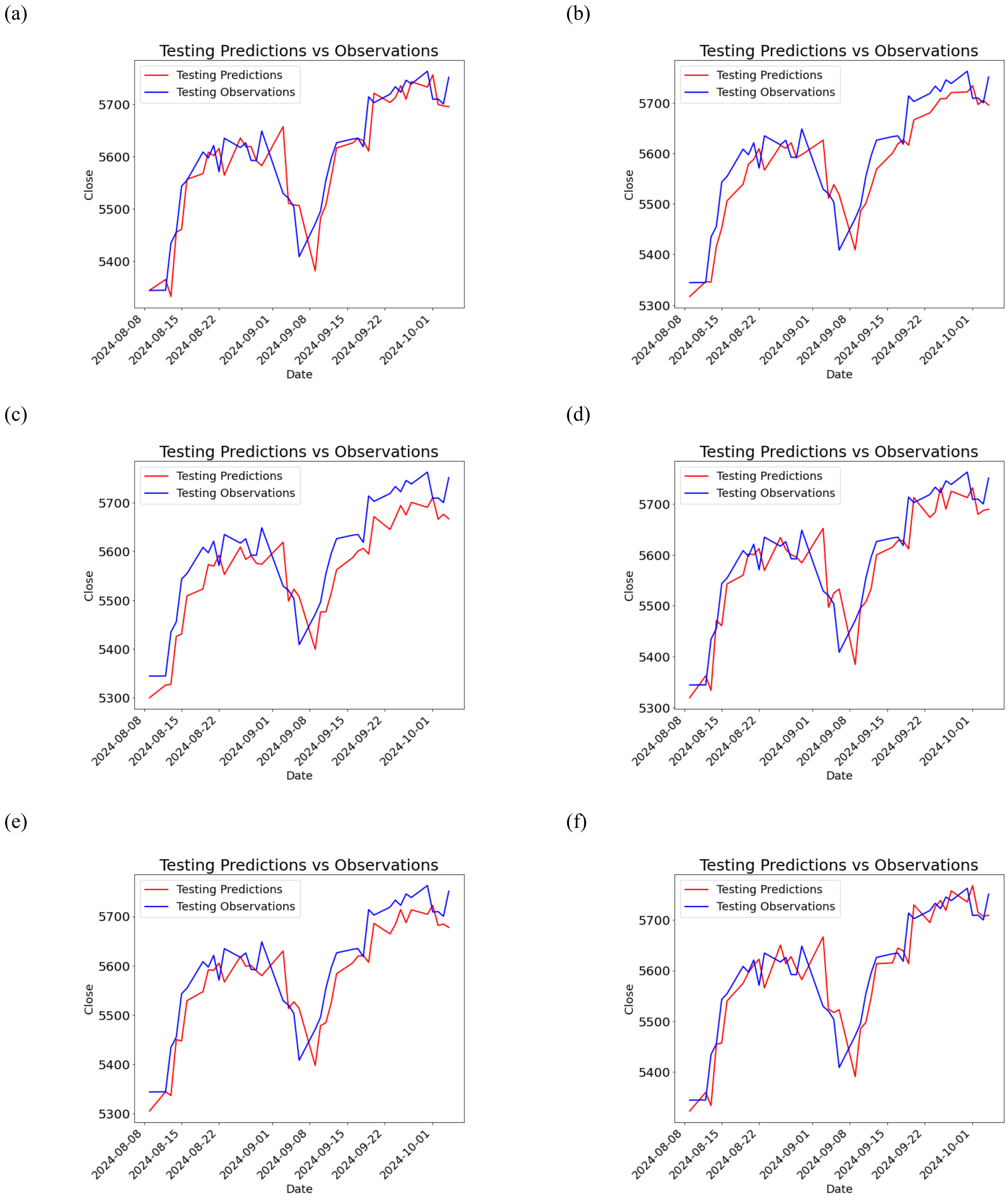

Fig. 4 Comparison of Predicted and Actual Results for the S&P 500 by (a) MACD-LSTM-ARIMA (proposed method), (b) LSTM, (c) GRU, (d) Bidirectional LSTM, (e) Support Vector Regression, (f) ARIMA.

Table I Mean Absolute Percentage Error (MAPE) of The Proposed Market Prediction Algorithms Versus Others Methods on Various Time-Series Datasets.

	Proposed Method	LSTM	Bi-LSTM	GRU	ARIMA	SVR
S&P500	**0.59%**	0.71%	0.70%	0.92%	0.65%	0.73%
TAIEX	**0.98%**	1.21%	1.09%	1.20%	0.99%	0.97%
NASDAQ	**0.91%**	0.89%	0.93%	0.96%	0.87%	0.92%
USDTWD	**0.46%**	0.48%	0.44%	0.42%	0.42%	0.45%
US10Y	**1.07%**	1.14%	1.11%	1.02%	1.06%	1.06%
GC	**0.61%**	1.25%	1.63%	2.74%	0.62%	1.91%
CCFC	**1.23%**	1.63%	1.38%	1.32%	1.39%	1.34%
Total	**0.836%**	1.044%	1.040%	1.226%	0.857%	1.054%

III. Experimental Result

In this section, we employ data from the *yfinance* database covering the period from March 5, 2023, to October 5, 2024. Ten percent of the time series observations are held out for testing, while the remaining 90 percent serve as the training set. The dataset includes the S&P500, the Taiwan Stock Exchange Weighted Index, the NASDAQ Composite Index, the USD/TWD exchange rate, the U.S. 10-year Treasury yield, gold futures, and continuous corn futures contracts. To assess the efficacy of our MACD–LSTM–ARIMA approach, we compare its forecasting performance against several established models—namely, a standard LSTM, a GRU, a bidirectional LSTM, support vector regression, and a standalone ARIMA model. Each model was selected based on relevant economic literature [11-15] and appropriately adjusted to suit our input data.

In this study, we use the mean absolute percentage error (MAPE) to quantify forecasting accuracy:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} |(y_t - \hat{y}_t)/y_t| * 100\%, \quad (7)$$

where y_t is the actual value at time t, $\hat{y}_t$ is the predicted value at time t, and n is the number of observations. A lower MAPE indicates better performance, and this metric enables aggregation for a comprehensive comparison.

Table displays the error metrics quantified by the mean absolute percentage error (MAPE), while Fig. illustrates—using the S&P 500 series as an example—a visual comparison of predicted versus observed values across different regression forecasting models.

As shown in Table , the proposed method achieves the lowest MAPE on the S&P 500, GC (gold futures), and CCFC (continuous corn futures) series. Moreover, when summing MAPE across all assets, the proposed method again ranks first, demonstrating strong adaptability to diverse datasets. Notably, in the GC row only the standalone ARIMA model attains a competitive MAPE of 0.62 %, whereas our hybrid approach—despite its LSTM component—yields an improved MAPE of 0.61%. We attribute this improvement to the explicit decomposition into high- and low-frequency components and the complementary strengths of the LSTM and ARIMA models, which together mitigate each other's weaknesses.

IV. Conclusions

The key feature of our approach is the integration of economic technical indicators with a long-memory machine-learning model (LSTM) and a classical short-term forecasting regression model (ARIMA). By introducing our MACD decomposition, one-dimensional time-series signals are decomposed via MACD calculations—implicitly reflecting how market participants rely on technical indicators when making trading decisions. Empirical results show that, compared to conventional models, our hybrid framework adapts more effectively across diverse univariate datasets and achieves superior predictive performance.

References

[1] A. A. A. Aguirre, R. A. R. Medina, and N. D. D. Méndez, "Machine learning applied in the stock market through the Moving Average Convergence Divergence (MACD) indicator," *Investment Management & Financial Innovations*, vol. 17, volume 4, article 44, 2020.

[2] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artificial Intelligence Review*, vol. 53, issue 8, pp. 5929-5955, 2020.

[3] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, John Wiley & Sons, 2015.

[4] S. Bhanja and A. Das, "Impact of data normalization on deep neural network for time series forecasting," *arXiv preprint arXiv*: 1812.05519, 2019.

[5] S. Arifin, A. K. Wijaya, R. Nariswari, and F. Suryanegara, “Long short-term memory (LSTM): Trends and future research potential,” *Int. J. Emerging Technology and Advanced Engineering*, vol. 13, issue 5, pp. 24-35, 2023.

[6] J. E. Cavanaugh and A. A. Neath, “The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 11, issue 3, article, e1460, 2019.

[7] H. M. Z. Raza, G. Haider, and S. Z. Haider, “Forecasting stock prices: Exploring the potential of ARIMA model for short-term predictions,” *Int. J. Management Research and Emerging Sciences*, vol. 14, issue 4, pp. 1–21, 2024.

[8] T. Zhang, S. Song, S. Li, L. Ma, S. Pan, and L. Han, “Research on gas concentration prediction models based on LSTM multidimensional time series,” *Energies*, vol. 12, issue 1, article 161, 2019.

[9] R. Porselvi and A. Meenakshi, “A study on the effectiveness of MACD,” *Educational Administration: Theory and Practice*, vol. 30, issue 5, 2024.

[10] R. M. West, “Best practice in statistics: The use of log transformation,” *Annals of Clinical Biochemistry*, vol. 59, issue 3, pp. 162–165, 2022.

[11] V. Gupta and A. Pandey, “Crude oil price prediction using LSTM networks,” *Int. J. Computer and Information Engineering*, vol. 12, issue 3, pp. 226–230, 2018.

[12] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using RNN encoder–decoder for statistical machine translation,” *arXiv preprint arXiv*: 1406.1078, 2014.

[13] Z. Cui, R. Ke, Z. Pu, and Y. Wang, “Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction,” *arXiv preprint arXiv*: 1801.02143, 2018.

[14] R. K. Dash, T. N. Nguyen, K. Cengiz, and A. Sharma, “Fine-tuned support vector regression model for stock predictions,” *Neural Computing and Applications*, vol. 35, issue 32, pp. 23295–23309, 2023.

[15] G. P. Zhang, “Time series forecasting using a hybrid ARIMA and neural network model,” *Neurocomputing*, vol. 50, pp. 159–175, 2003.