Journal of Economics and Public Finance
ISSN 2377-1038 (Print) ISSN 2377-1046 (Online)
Vol. 12, No. 1, 2026
www.scholink.org/ojs/index.php/jepf

Original Paper

Machine Learning-Based Prediction of Stock Market Returns

Yuqian Zhou[1*] & Liang Wang[2]

[1,2] Xizang University, Xizang, Lhasa, China

[*] Corresponding author, Yuqian Zhou

Received: January 6, 2026 Accepted: January 30, 2026 Online Published: February 2, 2026

doi:10.22158/jepf.v12n1p1 URL: http://dx.doi.org/10.22158/jepf.v12n1p1

Abstract

Focusing on the prediction of CSI 300 Index returns, this empirical study employs a hybrid CNN-LSTM-Attention model. The model integrates the strengths of CNN for local feature extraction, LSTM for temporal dependency modelling, and the Attention mechanism for key information focus, effectively capturing the multi-scale characteristics of financial data. Comparative experimental results demonstrate that multivariate models achieve superior fitting performance compared to univariate models, with the hybrid model outperforming either single model. This research validates the application value of deep learning models in financial time series forecasting, providing a novel approach for stock return prediction and offering reference for quantitative analysis and decision-making in financial markets.

Keywords

Machine learning, Stock market, Return prediction, CNN-LSTM-Attention model

1. Introduction

The stock market possesses the dual characteristics of high returns and high risk, impacting not only the wealth of individual investors but also the operational efficiency of the macroeconomic system. Consequently, research into predicting stock market returns has been a prominent topic of discussion among numerous scholars. From the 1960s onwards, numerous scholars began researching asset pricing models. Sharpe (1964), Lintner (1965), and Mossin (1966) proposed and refined the Capital Asset Pricing Model (CAPM), establishing for the first time a linear relationship between expected returns and risk. However, this model still suffers from issues such as overly simplistic assumptions and the explanatory power of the risk factor β being insufficient. Recognising the limitations of CAPM, scholars subsequently developed the Arbitrage Pricing Theory (APT) and the Fama-French multi-factor model. These breakthroughs transcended the constraints of single-factor models, significantly enriching the theoretical foundations of asset pricing. As financial theory continues to evolve, multiple novel characteristic factors have been introduced. For instance, Carhart (1977) proposed the momentum factor, creating a four-factor model; Fama and French (2015) incorporated earnings and investment characteristics into a three-factor model, validating the new framework's efficacy; Stambaugh and Yuan (2017) constructed a 'mispricing' four-factor model. Compared to models incorporating only market and size factors, this new framework more effectively adapts to market environments characterised by substantial anomalous phenomena. Traditional asset pricing models pioneered conventional methods for measuring returns, yet such models fundamentally rely on linear assumptions and structured data, rendering them ill-equipped to capture dynamic market shifts.

Compared to traditional linear regression methods, machine learning demonstrates superior efficacy in modelling financial return forecasting due to its robust non-linear analytical capabilities. Specifically, it enhances the accuracy of financial return prediction modelling by identifying complex interrelationships among different variables within the model. The literature on machine learning for stock market return forecasting can be broadly categorised into two types: traditional machine learning models centred on support vector machines (SVM) and decision trees. For instance, Fan, Pengying et al. (2025) focused on stock extreme value information, constructing the B-CARS model to demonstrate that models predicting stock return direction based on both highest and lowest prices significantly outperform single-extreme-value models or conventional models. The second category comprises deep learning models represented by neural networks. Deep learning, a pivotal branch within machine learning, enhances big data analysis capabilities through dimensionality reduction and feature extraction. Compared to traditional machine learning approaches, deep learning methods effectively process non-linear data, offering distinct advantages in studying market fluctuations. Consequently, it has increasingly become a focal point in contemporary financial research. Tsantekidis et al. (2017) found, through comparative analysis, that the CNN model holds a relative advantage in stock price forecasting. Li Peiran and Yang Lu (2023) observed that Transformer models can effectively identify market noise signals arising from lottery effect preferences and irrational trading. Further targeted empirical research grounded in behavioural finance theory revealed that market participants' lagging responsiveness to information is the primary factor generating momentum effects. With societal technological advancement, machine learning in stock return forecasting exhibits trends towards multi-model integration and expanded data dimensions. On one hand, diverse hybrid models continue to emerge, combining complementary machine learning algorithms to enhance learning capabilities. On the other, data sources have expanded beyond structured financial data to incorporate alternative unstructured data such as investor sentiment, financial news texts, and industry knowledge graphs, offering fresh perspectives for return forecasting research.

Despite significant advances in machine learning within financial research, where CNNs, LSTMs and Attention mechanisms have been applied to distinct financial domains respectively, each demonstrating strong performance advantages in their respective specialities and providing diverse technical support for quantitative finance studies, virtually no scholars have employed a model integrating all three for stock return prediction research. This paper aims to harness the synergistic potential of CNNs in capturing local patterns, LSTMs in handling long-term dependencies, and Attention mechanisms in dynamic feature selection. Building upon prior scholarly work, it innovatively applies a CNN-LSTM-Attention model to provide a more efficient technical pathway and theoretical foundation for stock return forecasting.

2. Theory and Models

2.1 Convolutional Neural Network

Convolutional neural networks (CNNs) represent a deep learning methodology frequently employed for image classification tasks. Structurally, a CNN primarily comprises fundamental components including an input layer, convolutional layers, pooling layers, fully connected layers, and an output layer, as illustrated in Figure 1. Within this framework, convolutional kernels continuously slide across the input data to extract local features, thereby enabling deep mining of feature information. Pooling layers typically follow convolutional layers, performing dimensionality reduction on the output features. Fully-connected layers then fuse and compress these features, ultimately outputting classification or regression results via activation functions.

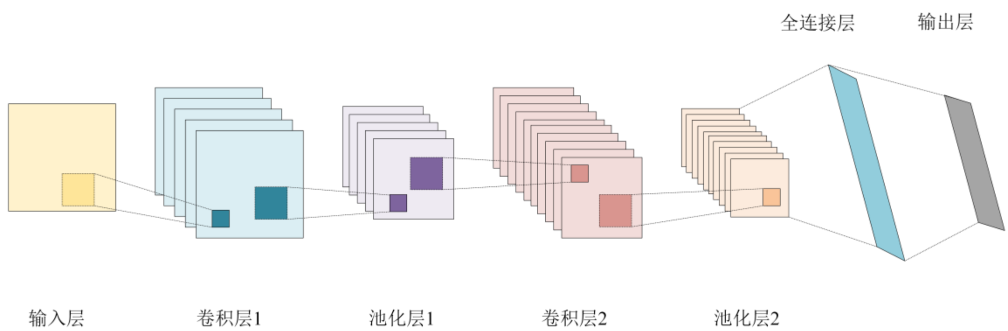

Figure 1. Schematic Diagram of CNN Architecture

2.2 Long Short-Term Memory

The LSTM model incorporates gating units and memory units within its internal architecture, building upon the foundation of RNN models. The gating units primarily regulate the input and output of information, while the memory units ensure that information maintains relevant states over time. Through this distinctive mechanism, LSTM can more efficiently retain or discard information when processing sequential data, thereby significantly enhancing the precision and reliability of capturing long-term dependencies and key insights within the data.

The internal structure of each unit within the LSTM model is primarily composed of input gates, forget gates, and output gates, with its cellular memory architecture illustrated in Figure 2.

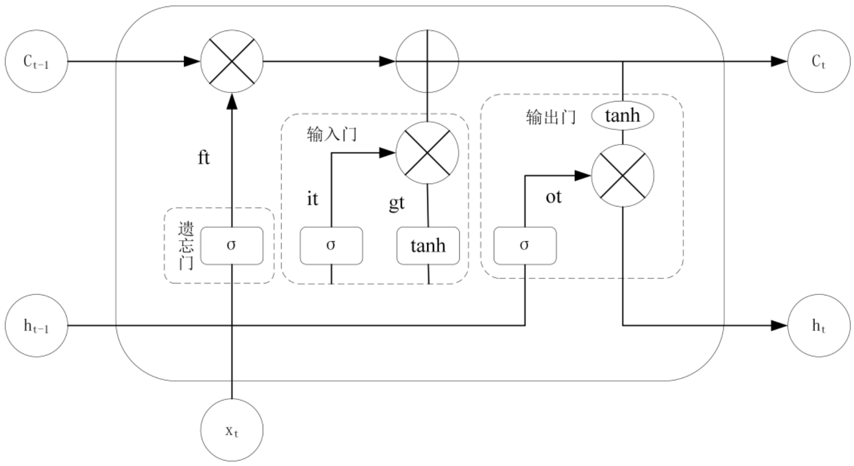

Figure 2. Schematic Diagram of the LSTM Architecture

(1) Forgetting Gate

The function of the forgetting gate is to filter out information that needs to be discarded. It achieves this filtering mechanism by activating the sigmoid function:

$$f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f) \tag{1}$$

(2) Input gate

The input gate of the LSTM is responsible for executing the mechanism of updating information, selectively retaining input content.

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \tag{2}$$

$$C_t = \tanh(W_c \times [h_{t-1}, x_t] + b_c) \tag{3}$$

$$C_t = f_t \times C_{t-1} + i_t \times C_t \tag{4}$$

(3) Output Gate

The output gate of the LSTM is responsible for controlling the output of cellular information. Its mathematical expression is as follows:

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t \times \tanh(C_t) \tag{6}$$

2.3 Attention Mechanism

Attention Mechanism (AM) enables networks to focus computational resources on a small amount of critical data by amplifying key components of input data while suppressing secondary information. Its primary structure comprises three vector spaces: Query (Q), Key (K), and Value (V). A schematic flowchart is illustrated in Figure 3.

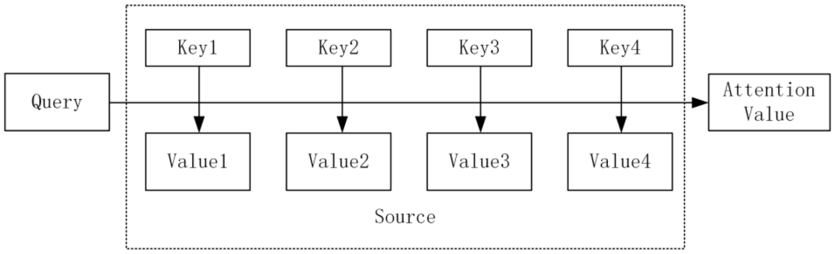

Figure 3. Schematic Diagram of Attention Mechanism Structure

2.4 CNN-LSTM-Attention Model

Compared to traditional models and single architectures, adopting a model fusion strategy effectively integrates the characteristics of different models, significantly enhancing the ability to learn and model the intrinsic patterns of time series data. The CNN-LSTM-Attention model developed in this study by organically combining these three approaches provides novel insights and methodologies for stock return prediction.

Within this model architecture, the CNN first processes the input stock market data to extract local features such as short-term price fluctuations and volume changes; LSTM then leverages these local features to further uncover underlying long-term dependency information, enabling deep learning and encoding of stock return trends; finally, the attention mechanism generates attention weight distributions through dot product operations. This dynamically identifies and filters genuinely valuable information from vast stock market datasets, better adapting to the complex and volatile nature of stock markets. It provides investors with more valuable references for decision-making.

3. Construction of a Stock Market Yield Prediction Model

3.1 Data Sources and Preprocessing

All data utilised in this study originates from the Wind database, encompassing daily trading data for the CSI 300 Index and its constituent stocks. The data spans the period from 5 January 2015 to 31 December 2024. To validate the model's predictive efficiency and feasibility, the dataset was divided proportionally into an 8:2 split, comprising the training set and test set respectively. To ensure the model's predictive capabilities met expectations, the pandas_ta module—an open-source library within Python specifically designed for financial technical indicator calculations—was incorporated. This enabled the generation of multiple technical indicators from the raw data.

Table 1. Indicators and Their Definitions

name	meaning	name	meaning
open	Opening price	RET	Yield Indicator
close	Closing price	BIAS_SMA_6	6-day Deviation Rate
high	Highest price	RSI_14	14-day Relative Strength Index
low	Lowest price	ROC_12	12-day Rate of Change Indicator
volume	Transaction volume	MOM_10	10-day Momentum Indicator
money	Transaction value	K	Stochastic Oscillator K-line
SMA_5	5-day simple moving average	D	Stochastic Oscillator D-line
SMA_10	10-day simple moving average	J	Stochastic Oscillator J-line
SMA_15	15-day simple moving average	Lower	Upper Bollinger Band
SMA_60	60-day simple moving average	Middle	Middle Bollinger Band
SMA_120	120-day simple moving average	Upper	Lower Bollinger Band

To eliminate the influence of data dimensions, the dataset undergoes min-max scaling prior to model fitting, with the specific procedure as follows:

$$x_{scaled} = \frac{x - x_{\min}}{x_{\max} - x_{\max}} \quad (7)$$

x_{min} and x_{max} representing the minimum and maximum values of the feature respectively, x_{scaled} For the scaled value.

3.2 Feature Selection

It is generally accepted that the explanatory power of asset pricing models regarding market information correlates positively with the number of variables incorporated, meaning that including more market information variables enhances predictive accuracy. However, existing research indicates that high-dimensional data often proves incompatible with traditional methods such as linear assumptions. Noise interference and feature redundancy within the data intensify as the number of variables increases, potentially leading to overfitting and significant deviations in model predictions. Jiang Fuwei et al. (2022) conducted empirical research on China's A-share market, demonstrating that while multi-factor models incorporating excessive variables may fit data within the sample, they exhibit significantly increased errors in out-of-sample predictions due to over-capturing noise. This validates the incompatibility between high-dimensional data and traditional linear methods. To balance comprehensive information coverage with model parsimony while maximising predictive efficiency, this study employs Pearson correlation analysis and random forest methods to scientifically and effectively screen existing feature variables. The former adopts a linear perspective, identifying redundant features exhibiting high linear dependence by calculating correlation coefficients between variables. Random forests, conversely, refine the feature set from a non-linear viewpoint. Based on decision tree ensemble principles, they quantitatively analyse each feature's contribution to model predictions. Their advantage lies in handling complex non-linear relationships between variables and effectively identifying features with critical influence on the prediction target.

Random forests can uncover latent associative features elusive to linear analysis, complementing Pearson correlation analysis to ensure the filtered feature subset comprehensively covers relevant information while maintaining high predictive efficiency. This approach maximises the retention of variables exhibiting significant explanatory power for stock return forecasting.

3.2.1 Pearson Correlation Analysis

Pearson correlation analysis is frequently employed to measure the strength and direction of linear relationships between two continuous variables. It constitutes a standardised processing method that utilises the ratio of covariance to variable standard deviation to eliminate the influence of variable dimensions, thereby objectively reflecting the degree of linear association between variables. This study employs heatmap visualisation techniques to intuitively present the linear correlation between features, as illustrated in Figure 4. The intensity of colours in the figure exhibits a strong correlation with the magnitude of the absolute correlation coefficient. For instance, warm-toned areas (red spectrum) denote the strength of positive correlations, while cool-toned areas (blue spectrum) reflect the degree of negative correlations. The distribution observed in the correlation heatmap indicates that most features exhibit varying degrees of linear dependence.

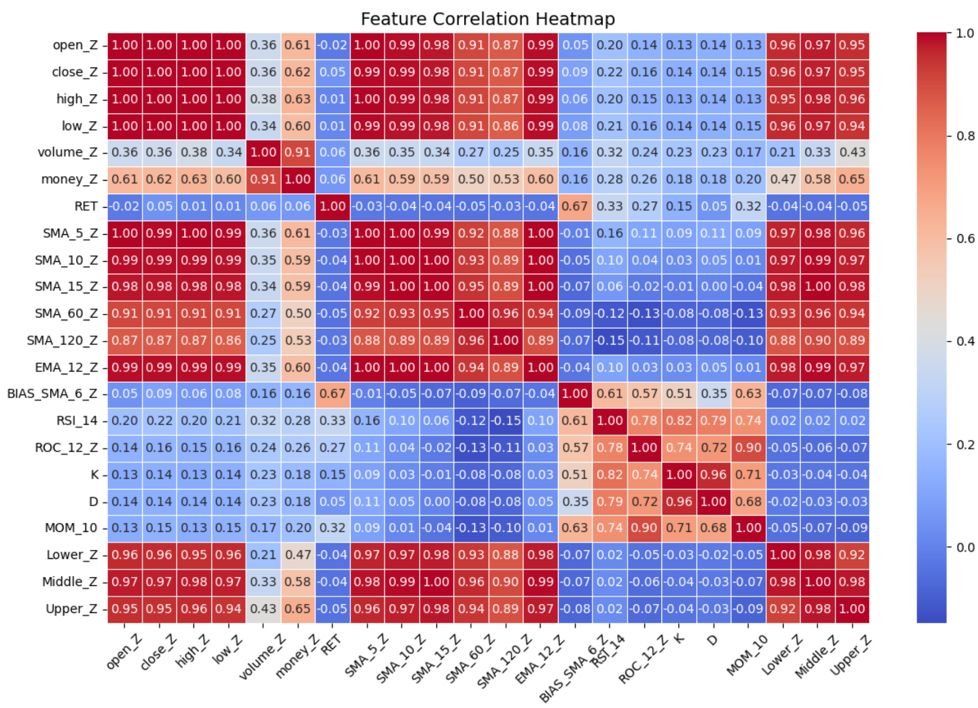

Figure 4. Results of Pearson Correlation Analysis

As illustrated in the figure, price-related indicators such as the opening price (open) and closing price (close) exhibit multiple linear relationships with other features. However, as these price characteristics directly record the raw state of market transactions and contain substantial latent information, it has been decided to retain them in their entirety. Volume indicators (volume_Z) exhibit a strong positive correlation with capital flow indicators (money_Z), signifying that capital dynamics directly influence market trading activity. Moving average indicators, exemplified by SMA, demonstrate not only high internal positive correlation but also strong associations with price-based indicators. This aligns with the nature of moving averages, which are calculated from price sequences to smooth and present price trends, enabling effective tracking of price movement changes. Technical indices such as RSI_14 exhibit moderate positive correlations both among themselves and with certain moving average indicators. Such indicators demonstrate synergistic capabilities in reflecting market overbought/oversold conditions and momentum shifts, thereby aiding trend analysis. Bollinger Band-related metrics like Lower_Z and Middle_Z can delineate price fluctuation ranges, providing reference points for assessing price boundaries. In summary, an initial screening of indicators is conducted.

3.2.2 Random Forest

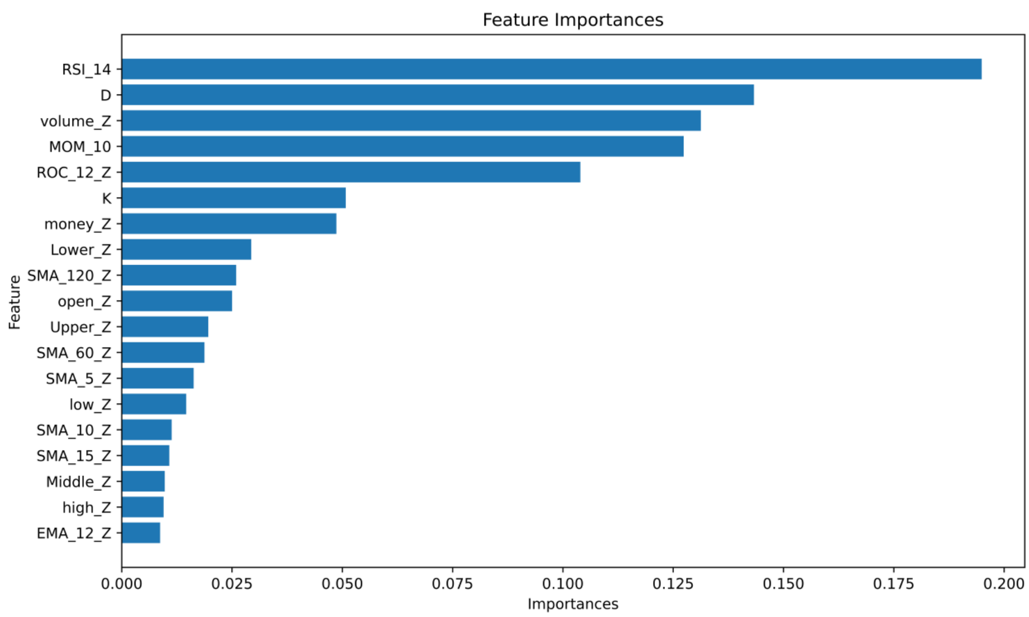

Figure 5. Random Forest Feature Selection Results

The feature importance chart derived from the random forest reveals significant variations in the importance of different features to the model. Among these, RSI_14 exhibits the highest feature importance, indicating that during the random forest model's learning process, the 14-period Relative Strength Index data contributes most significantly to the prediction outcomes. This indicator effectively reflects the comparison of buying and selling forces within the short-term market and the relative strength of price movements, providing the model with crucial market sentiment information. Following closely are variables such as D and MOM_10, demonstrating their considerable importance within the model by providing valuable predictive insights from distinct perspectives; Features such as EMA_12_Z and high_Z exhibit relatively lower importance, suggesting their weaker influence on prediction outcomes within the random forest model. This may stem from their information being partially covered by other features, or their unique value failing to be fully demonstrated within the current dataset and prediction objectives.

This study constructed a multidimensional feature selection framework through cross-validation of Pearson correlation heatmaps and random forest importance assessments. By integrating results from correlation analysis and random forest dual screening, the research simultaneously avoided information redundancy and multicollinearity issues inherent in high-dimensional data while ensuring the feature set comprehensively covered critical market information. The selected feature set comprises open, close, high, low, BIAS_SMA_6, RSI_14, ROC_12, K, D, money, lower, upper, and SMA_5. This approach maximises market information coverage while minimising feature redundancy, thereby enhancing the model's effectiveness and accuracy in predicting stock returns.

3.3 Evaluation Indicators

This study drew upon prior scholarly research when constructing the comprehensive evaluation metric system, selecting RMSE, MAE, R^2, and MAPE as core assessment indicators to systematically evaluate model predictive performance across multiple dimensions. RMSE and MAE focus on measuring absolute error, while MAPE examines relative error performance. R^2 provides an overall assessment of the model's fit. This evaluation framework enables a comprehensive and objective assessment of model performance, offering a scientific basis for optimisation, refinement, and practical application.

3.4 Parameter Selection

The software employed in this research and its version is Python 3.9, utilising the following modules: Numpy 2.2.4, Pandas 2.2.3, pandas_ta 0.3.14b, and matplotlib 3.10.1. Additionally, the specific parameters configured for the CNN-LSTM-Attention model are as follows:

Table 2. Specific Parameters of the CNN-LSTM-Attention Model

	Parameters	Parameter value
CNN layer	Number of convolutional kernels	64
	Convolutional kernel size	3*3
	Activation function	ReLU
	padding	1
LSTM layer	Input feature dimension	64
	Number of hidden units	50
	Activation function	Sigmoid
Attention layer	Activation function	Softmax
Training configuration parameters	Output dimension	1

	Optimizer	Adam
	Learning rate	0.001
	Number of training epochs	200
	Loss function	MSELoss
	Batch_size	128

4. Analysis of Stock Market Yield Forecast Model Results

4.1 Single-Variable Model Fitting Results

In the single-variable forecasting scenario utilising only historical data from the CSI 300 Index as input features, the fitting performance of each model exhibited significant variation. In terms of error metrics, the CNN-LSTM-Attention model—which integrates a convolutional neural network (CNN) with an attention mechanism—delivered the most favourable performance. Its root mean square error (RMSE) was as low as 53.4049, while its mean absolute error (MAE) stood at 37.7155, representing the smallest values among all models. Concurrently, this model achieved a coefficient of determination (R^2) of 0.9619, approaching unity, indicating its strongest interpretability of data trends. Its mean absolute percentage error (MAPE) of merely 1.03% further validated the reliability of its predictive accuracy. By contrast, the baseline LSTM model performed worst, with an RMSE of 81.9025 and R^2 of 0.9105. The disparity with the optimal model demonstrates the effectiveness of incorporating CNN feature extraction capabilities and the Attention weight allocation mechanism in enhancing univariate time series forecasting performance. Notably, the LSTM-Attention model—which merely superimposes the Attention mechanism (RMSE=74.5581, R^2=0.9258)—yields inferior results to the CNN-LSTM hybrid model (RMSE=57.7655, R^2=0.9555). This indicates that in univariate scenarios, convolutional extraction of local features may yield greater predictive improvements than weight optimisation via Attention mechanisms.

Table 3. Comparison of Single-Variable Model Fitting Results

Model Name	RMSE	MAE	R2	MAPE
LSTM	81.9025	54.5159	0.9105	1.47%
CNN-LSTM	57.7655	46.4615	0.9555	1.23%
LSTM-Attention	74.5581	55.6849	0.9258	1.52%
CNN-LSTM-Attention	53.4049	37.7155	0.9619	1.03%

4.2 Multivariate Model Fitting Results

When incorporating multi-dimensional relevant features (open, close, high, low, BIAS_SMA_6, RSI_14, ROC_12, K, D, money, lower, upper, SMA_5) for prediction, the overall performance of each model improved compared to the univariate scenario, and the ranking of model performance changed. Among these, the CNN-LSTM-Attention model maintained its advantage, with RMSE reduced to 47.0523, R^2

improved to 0.9705, and MAPE at merely 0.86%. This demonstrates the model's enhanced capability to fuse complex features under multivariate input conditions. Notably, the LSTM-Attention model demonstrated significant improvement in the multi-variable scenario, with its RMSE decreasing from 74.5581 in the single-variable scenario to 57.5020, MAE at 34.0478, and MAPE at 0.91%, placing it second only to the optimal model. This indicates that the attention mechanism can more accurately capture the temporal correlations of key influencing factors when processing multi-source input features. Moreover, the base LSTM model's R^2 in the multi-variable scenario increased from 0.9105 to 0.9526, with substantial optimisation of error metrics. This validates the effectiveness of incorporating external correlation features in enhancing the predictive accuracy of the CSI 300 Index. Conversely, the CNN-LSTM model exhibited largely consistent performance across single- and multi-variable scenarios, potentially due to the limited efficiency of its convolutional layers in extracting multi-variable features.

Table 4. Comparison of Multivariate Model Fitting Results

Model Name	RMSE	MAE	R2	MAPE
LSTM	59.5908	40.2923	0.9526	1.08%
CNN-LSTM	57.7655	46.4615	0.9555	1.25%
LSTM-Attention	57.5020	34.0478	0.9566	0.91%
CNN-LSTM-Attention	47.0523	31.9029	0.9705	0.86%

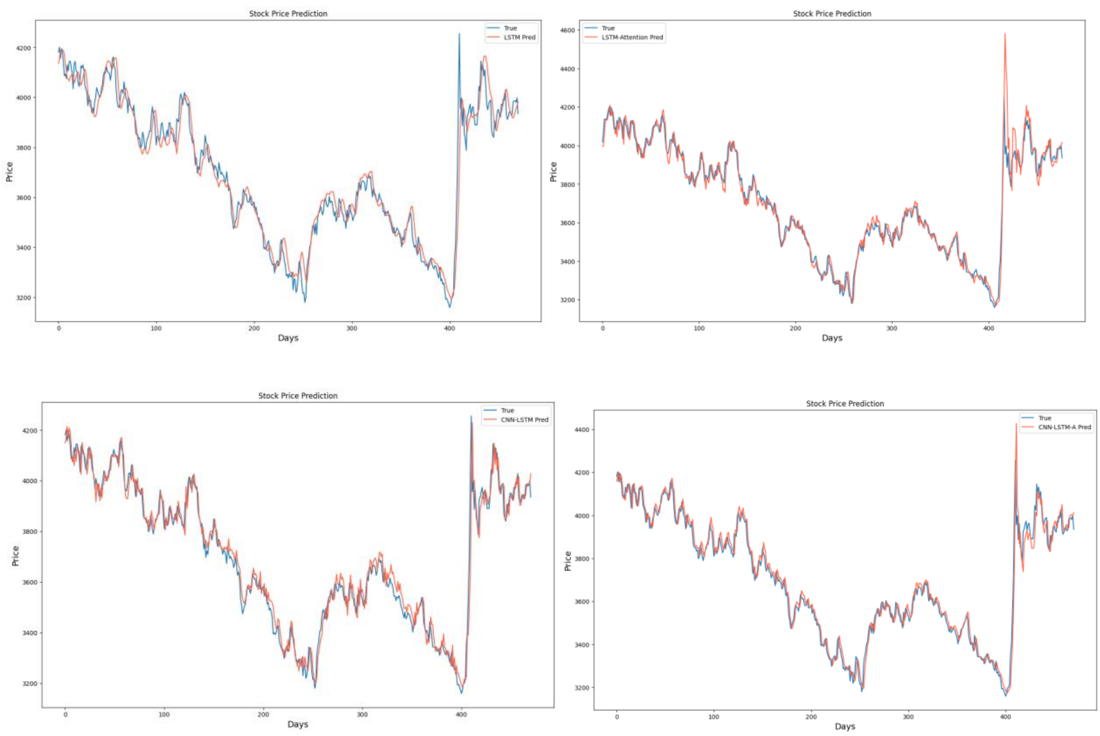

Figure 6. Multivariate Model Fitting Results

5. Discussion

This study employs a CNN-LSTM-Attention model to forecast returns on the CSI 300 Index, marking a pivotal advancement in financial time series analysis. By employing Pearson correlation heatmaps to dissect the interrelationships among financial indicators, it elucidates the intrinsic connections between price, volume, and technical indicators, thereby establishing a robust foundation for feature selection in model inputs. Fitting results indicate: (1) Multivariate models more comprehensively reflect market dynamics, encompassing richer information, thus yielding superior fitting outcomes compared to univariate models. (2) Hybrid models overcome the limitations of single-model approaches through structural synergy, achieving more precise fitting of stock return variations.

Overall, this study validates the application potential of deep learning models in financial return forecasting while revealing the challenges posed by market complexity to precise prediction. It aims to provide insights for quantitative financial investment and risk management practices, driving deeper integration of artificial intelligence and financial analysis within fintech. This will empower market participants to respond to volatility more rationally and scientifically, advancing the ongoing exploration of financial market patterns.

Fund Project

Xizang University Postgraduate High-Level Talent Cultivation Programme Project: 'Machine Learning-Based Prediction of Stock Market Returns' (2025-GSP-S008).

References

Alzubaidi, L., Zhang, J., Humaidi, A. I., et al. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*, 1-74. https://doi.org/10.1186/s40537-021-00444-8

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57-82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Fama, E. F., & French, K. R. (1993). Common Risk Factors in Returns on Stocks and Bonds. *Journal of Financial Economics*, *33*(1), 3-56. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1-22. https://doi.org/10.1016/j.jfineco.2014.10.010

Fang Yi, Chen Yuzhi, & Wei Jian. (2022). Artificial intelligence and the Chinese stock market: A quantitative portfolio study based on machine learning predictions. *Industrial Technology Economics*, *41*(08), 83-91. (in Chinese)

Fan, P. Y., Han, J. C., Xie, H. B., et al. (2025). Directional Prediction of Stock Returns: Two Poles Are Better Than One – An Empirical Analysis Based on the B-CARS Model. *Systems Engineering Theory and Practice*, *45*(01), 93-108. (in Chinese)

Fang Siran, Guo Mingjun, & Wei Yunjie. (2023). Can Incorporating Public Sentiment Information Effectively Improve Exchange Rate Forecasting Accuracy? *Journal of Econometrics*, *3*(02), 464-486. (in Chinese)

Jiang Fuwei, Xue Hao, & Zhou Ming. (2022). Does big data enhance the pricing capability of multi-factor models? — An Exploration of China's A-Share Market Using Machine Learning Methods. *Systems Engineering Theory and Practice*, *42*(08), 2037-2048. (in Chinese)

Li Peiran, & Yang Lu. (2023). Can Deep Learning Resurrect Failed Factors? — An Exploration of Monthly Momentum Effects in China's A-Share Market Using Transformer Models. *Investment Research*, *42*(07), 22-46. (in Chinese)

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *Journal of Finance*, *20*(04), 587-615. https://doi.org/10.1111/j.1540-6261.1965.tb02930.x

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, *34*(04), 768-783. https://doi.org/10.2307/1910098

Mostafavi, M. S., & Hooman, R. A. (2025). Key technical indicators for stock market prediction. *Machine Learning with Applications*, *2025*, 20100631-100631. https://doi.org/10.1016/j.mlwa.2025.100631

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, *1976*, 341-360. https://doi.org/10.1016/0022-0531(76)90046-6

Sharpe, W. F. (1964). Capital Asset Prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, *19*(3), 425-442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Stambaugh, R. F., & Yuan, Y. (2017). Mispricing factors. *Review of Financial Studies*, *30*(4), 1270-1315. https://doi.org/10.1093/rfs/hhw107

Shi, S., Li, F., & Li, W. (2025). A Hybrid Long Short-Term Memory-Graph Convolutional Network Model for Enhanced Stock Return Prediction: Integrating Temporal and Spatial Dependencies. *Mathematics*, *13*(7), 1142-1142. https://doi.org/10.3390/math13071142

Tian Xianzhong, Hu Anna, & Gu Siyi. (2021). A stock market forecasting method based on time series chains. *Journal of Zhejiang University of Technology*, *49*(05), 503-510+563. (in Chinese)

Tsantekidis, A., Passalis, N., Tefas, A., et al. (2017). Forecasting Stock Prices from the Limit Order Book Using Convolutional Neural Networks. *IEEE*, 2017. https://doi.org/10.1109/CBI.2017.23

Tang Zhuoyuan. (2024). Green Bond Market Analysis and Index Forecasting: An Exploration Based on Knowledge Graphs and Deep Learning Technologies. *Business Exhibition Economy*, *2024*(18), 113-117. (in Chinese)

Xu Xuechen. (2024). Empowering the Financial Industry in the Digital Age with Large Language Models like ChatGPT: Privacy Protection, Algorithmic Discrimination and Systemic Risk. *Journal of Jinan University* (Philosophy and Social Sciences Edition), *46*(08), 108-122. (in Chinese)

Yin Haiyuan, & Wang Xiaoxiao. (2024). Does Investor Sentiment Affect Stock Pricing Efficiency? Textual Evidence from Stock Communities. *Journal of Beijing Union University* (Humanities and

Social Sciences Edition), *22*(03), 96-111. (in Chinese)

Yang, L., Wu, Y. Q., Wang, J. L., et al. (2018). Review of Recurrent Neural Networks. *Computer Applications*, *38*(S2), 1-6+26. (in Chinese)

Zhao, Y., Guo, Y., & Wang, X. (2025). Hybrid LSTM–Transformer Architecture with Multi-Scale Feature Fusion for High-Accuracy Gold Futures Price Forecasting. *Mathematics*, *13*(10), 1551-1551. https://doi.org/10.3390/math13101551

Zhou Jiani, Liu Chunyu, & Liu Jiapeng. (2025). Stock Price Trend Forecasting Model Integrating Channels and Multi-Head Attention. *Computer Engineering and Applications*, *61*(08), 324-338. (in Chinese)