Proceedings of CONF-MLA 2026 Symposium: Intelligent Systems and Automation: AI, IoT, Robotic Engineering & Algorithm
DOI: 10.54254/2755-2721/2026.LD35447

A Review of LSTM-Based Stock Prediction for Long-Short and Market-Neutral Portfolio Construction

Siqian Li

College of Science and Engineering, James Cook University Singapore, Singapore, Singapore
siqian.li@my.jcu.edu.au

Abstract. In recent years, deep learning algorithm models have gradually been applied to fields such as stock prediction, asset pricing, and portfolio management. Among them, Long Short-Term Memory (LSTM) has been widely used to predict stock prices, returns, and directional movements due to its ability to handle long-term dependencies in time series. Compared to traditional linear financial models, LSTM can better capture non-linear correlations and dynamic fluctuations embedded in financial data, which explains its significant attention in financial forecasting research. However, from a practical investment perspective, a lower prediction error does not necessarily translate into better portfolio performance. This paper reviews the application of LSTM in stock prediction, long-short strategies, and market-neutral portfolio construction. We first introduce the theoretical foundations of LSTM, alpha signals, long-short strategies, and market-neutral portfolios. Subsequently, it delineates the complete research pipeline spanning stock price forecasting, stock alpha ranking and portfolio construction, before analyzing the practical value, existing limitations and prospective research directions of LSTM-based frameworks. Core Viewpoint: The optimal positioning of LSTM is not to predict absolute stock prices, but to serve as an alpha signal generation tool for stock ranking and portfolio optimization. Future research should place more emphasis on point-in-time data, out-of-sample validation, risk neutralization, transaction cost modeling, and comprehensive investment workflow design to improve the usability of LSTM strategies in real-world trading environments.

Keywords: LSTM, stock prediction, long-short strategy, market-neutral portfolio

1. Introduction

Traditional financial asset pricing frameworks explain stock returns mainly as risk compensation mechanisms, usually through risk factors including market beta, size, value and momentum. Machine learning and deep learning models take a different route. They try to learn non-linear patterns from historical prices, trading volumes, fundamental data, and the wider market environment. As financial datasets have grown and computing power has improved, deep learning has been used more often in stock prediction, algorithmic trading, risk management, and portfolio management [1, 2].

Among deep learning models, Long Short-Term Memory (LSTM) networks are used frequently in financial time-series forecasting because they handle sequential data well. Stock prices and

DOI: 10.54254/2755-2721/2026.LD35447

returns are usually affected by past market conditions, and current price changes may also contain information such as investor behavior, market sentiment and liquidity changes. On this basis, LSTM models can learn time-varying patterns in historical data and use them to forecast future prices, excess returns and price directional trends [3].

Nevertheless, stock return forecasting and portfolio construction are not the same task. Many studies judge model performance with Mean Squared Error (MSE), Mean Absolute Error (MAE), or directional accuracy. These metrics can show statistical forecasting performance, but they do not directly show whether a model has investment value. For investors, what matters more is whether the model can boost annualized returns and the Sharpe ratio, reduce maximum drawdowns, and remain effective after accounting for transaction costs and risk exposures.

The essence of a long-short strategy is not to say how much you think the price of one stock will go up or down in absolute terms, but rather which stocks will do relatively better or worse going forward. Investors can buy the top ranked stocks and short the lowest ranked stocks to capture the return spread between the long and short sides. Market-neutral strategies take this one step further and require the portfolio to have minimal exposure to the overall market, so that the returns are derived more from the stock selection skill of the manager than from market beta. So, in the context of this study, LSTM is more to be seen in terms of an alpha signal generation technique or a return prediction model than a pure price prediction model.

This review focuses on how LSTM prediction results serve the mechanism of long-short and market-neutral portfolio construction. The paper will provide a comprehensive review across five areas: theoretical foundations, current research progress, application value, existing problems, and future directions, with an emphasis on the relationship between predictive ability and actual portfolio performance.

2. Theoretical foundation of LSTM-based portfolio strategies

LSTM was initially introduced by Hochreiter and Schmidhuber to address the gradient vanishing problem common in traditional recurrent neural networks during long-sequence learning [3]. By controlling information retention, updating, and output through input, forget, and output gates, LSTM can preserve useful information over long periods. This makes it very suitable for financial time series data, such as stock prices, volumes of stock trades, and technical indicators.

In stock prediction, LSTM usually uses the historical prices, trading volumes, stock returns and technical indicators as input of the model to predict future prices, expected returns and price direction trends. In portfolio management, however, prediction per se is not the end goal. The main issue is whether the predicted outputs from the model can be converted into tradable alpha signals. For example, stocks with highest predicted future return are given the highest rank and stocks with lowest predicted future returns are given the lowest rank. Through this approach, the output of LSTM transforms from a simple forecasting result into a portfolio signal.

Long-short strategies give a basic use case for this transformation. A long-short strategy in general is a stock-picking strategy that takes long positions in stocks that are expected to increase in value and short positions in stocks that are expected to decline in value, with returns derived primarily from the relative performance of the long and short portfolios. Therefore, the long-short strategy pays more attention to the relative strength across different stocks at a single moment in time, that is, cross-sectional ranking, rather than the time-series prediction single-stock price forecasting. That is, these strategies do not require accurate forecasts of future individual stock prices but only the ability to correctly identify, with some given confidence, in the model's ranking, stocks that will outperform in the future from those that would under perform. Market-neutral

DOI: 10.54254/2755-2721/2026.LD35447

strategies are an extension of long-short strategies. Although long-short equity strategies maintain both long and short positions, they are not truly market-neutral by construction. If the long portfolio has a high overall market beta exposure and the short side has a low beta exposure, the total portfolio will still be sensitive to broad market movements. Market-neutral portfolios need to be adjusted for market beta, sector, and style factor exposures (at the portfolio level) so that returns come more from the model's ability to pick stocks than from systematic risk.

The Fama-French three-factor and Carharts four-factor models illustrate that firm-return variation can be explained to a large extent by common factors such as market, size, value, and momentum [4, 5]. Market-neutral strategies seek to reduce these systemic risk exposures so that the portfolio returns are largely dependent on the security selection skill [6]. This offers a critical insight for asset portfolio strategies based on LSTM predictions: if an LSTM strategy does not control for these risk factors, its apparent alpha might just be returns from traditional factors. Consequently, the evaluation of LSTM models should extend beyond simple forecasting error metrics to further assess cross-sectional ranking power, realized portfolio returns, factor risk exposures and net performance after transaction costs.

3. Research status: from stock prediction to portfolio construction

The literature can be roughly divided into three stages of the sequence: the first stage— predicting stock price or direction; the second stage begins to apply LSTM to cross-sectional ranking and long-short strategies; the third stage incorporates LSTM signals further into market-neutral portfolio building.

3.1. LSTM for stock price and direction prediction

Early studies typically used OHLCV data (open, high, low, close prices, and trading volume) to train LSTM models to predict future stock prices or directional movements. Studies in this stage mainly aimed to benchmark the forecasting performance of LSTM against conventional machine learning algorithms, statistical time-series models and vanilla neural networks. Fischer and Krauss observe that the superiority of LSTM decayed substantially when transaction costs were incorporated [7]. The main contribution of their work was that it was not limited to simple statistical comparisons of forecast accuracy but also studied whether model forecasts could be converted into profitable trading strategies. They found that although LSTM attained good performance during certain intervals, its strategic superiority diminished significantly after transaction costs, and its excess returns were particularly unstable after 2010. This result shows that although LSTM can capture certain part of market predictive signals, such signals are limited in long-run stability and cannot be entirely translated into real-trading excess return. This initial phase advanced the application of LSTM in financial forecasting, but its limitations are quite obvious. Numerous studies overemphasized forecasting error metrics while neglecting in-depth analysis of how model predictive outputs can be integrated into portfolio construction workflows. In fact, more accurate price predictions do not automatically guarantee stronger stock-picking abilities or higher portfolio returns.

3.2. LSTM for cross-sectional ranking and long-short strategy

As the research has evolved, a few researchers have started to migrate from forecasting individual stocks to the cross-sectional ranking. For traditional long-short trading strategies, the main goal is

DOI: 10.54254/2755-2721/2026.LD35447

not to predict the absolute future price of an individual stock, but rather to find stocks that have greater expected relative outperformance at a given moment in time. That is, the model has to produce signals which can be used for ranking stocks.

The LSTM-DNN hybrid framework proposed by Hou et al. reflects this direction [8]. This model applies a LSTM to a sequence of historical returns, a DNN to a set of fundamental and market features, and finally a DNN that predicts the future returns, based on which a Top/Bottom long-short portfolio is constructed. The main novelty of this paper is the information combination at two levels, transforming LSTM from a price prediction model to a stock ranking and portfolio construction model. In this sense, within this setting, we can interpret the return predictions issued by LSTM as ranking signals. Investors go long (hold) the stocks with the best predicted returns and go short (sell) the bottom ones. If the model's ranking ability is effective, the long portfolio should outperform the short portfolio in the future. Unlike simple comparisons of statistical metrics such as MSE and MAE, this evaluation standard is far more consistent with real world investment practice.

3.3. LSTM for market-neutral portfolio construction

Recent studies have further embedded LSTM predictions into market-neutral portfolio construction. Compared to ordinary long-short strategies, market-neutral strategies not only require holding both long and short positions but also demand controlling market risk exposures so that portfolio returns remain largely independent of overall market shifts.

Cipiloglu Yildiz and Yildiz used LSTM-predicted returns to calculate portfolio weights, showing that LSTM predictions can improve portfolio performance [9]. However, their research mainly centered on long-only portfolios and failed to construct strictly market-neutral investment vehicles.The research conducted by Nafia et al. is more aligned with the core focus of this review. They used LSTM to predict the returns of S&P 500 consumer staples stocks and ranked them to construct a weekly rebalanced equity-market-neutral portfolio [10]. Their study formed a complete pipeline of "prediction - ranking - portfolio construction - performance evaluation" and evaluated the portfolio using metrics such as annualized return, Sharpe ratio, maximum drawdown, beta, and market correlation.

Overall, relevant research has gradually moved from pure stock price prediction to long-short strategies and market-neutral portfolio construction. This shift indicates that the value of LSTM in finance is moving from prediction accuracy to portfolio effectiveness.

4. Application value of LSTM in long-short and market-neutral portfolio construction

The first value of LSTM in long-short and market-neutral portfolios is its potential to generate more effective cross-sectional ranking signals than traditional linear models. Traditional multi-factor linear models usually explain or predict stock returns through factors like market, size, value, and momentum. While highly interpretable, these models rely on relatively stable linear relationships and struggle to fully capture non-linear shifts, temporal dynamics, and variable interactions in stock returns. In contrast, LSTM can learn dynamic patterns from sequential data such as historical prices, returns, volumes, and volatility. Therefore, compared with conventional linear financial models, LSTM holds a comparative edge in cross-sectional equity ranking, as it jointly captures time-varying effects, nonlinear correlations and multivariate factor interactions. If further combined with fundamental, industry, and factor variables, LSTM can distinguish future relative winners and losers more effectively. For long-short and market-neutral strategies, portfolio returns mainly come from the performance gap between long and short stocks, making ranking capability far more critical than

DOI: 10.54254/2755-2721/2026.LD35447

simply reducing prediction errors. Based on these ranking results, an LSTM-based framework can translate predicted returns into long portfolios, short portfolios, weight allocations, or trading signals, moving the model from the forecasting layer into the portfolio construction layer. Consequently, model evaluation is no longer just about whether the prediction error is lower, but whether the predictive signal truly improves portfolio performance.

The second benefit is in giving a cleaner way to test for alpha in a market neutral setting. Returns of regular long-only approaches can be heavily skewed by broad market rallies and it's difficult to tell if a return is driven by ability of a model or by market beta. Market-neutral strategies allow for more robust confirmation of a model's idiosyncratic stock-selection alpha by retaining paired long-short exposures and by limiting market and style factor risks at the portfolio level. If an LSTM strategy can generate stable returns even after controlling for market risk, factor exposures, and transaction costs, it is highly likely to provide a genuinely effective alpha signal. In short, the market-neutral framework allows researchers to distinguish genuine stock-selection value generated by the model from superficial returns driven by unconstrained systematic risk exposures. The third value is its ability to complement traditional asset pricing theories. Conventional asset pricing and factor models come with interpretable financial intuition whereas LSTM models are powerful at modeling complex nonlinear dependencies and time-varying return dynamics. Hence, LSTM need not replace traditional factor models; rather, it can be used with the Fama-French factors, momentum factors, industry dummies, fundamental dummies, and dummies for macroeconomic conditions to model more complex mechanisms of generating returns. Related literature also suggests that much of the ML's improvement in predicting stock returns comes from its ability to model nonlinear relations and interactions among variables [11]. Should deep learning models be imbued with asset pricing constraints and financial theories, this forecast could be a lot more believable than pure black-box models [12]. Hence when constructing long-short and market-neutral portfolios, the appropriate place of LSTM is not to completely replace traditional financial models, but to serve as predictive instrument that improves ranking, stock-selection, and portfolio construction performance.

5. Main problems in existing research

First, overfitting constitutes a prominent flaw in current studies. Financial datasets are inherently noisy, market regimes evolve dynamically, and LSTM architectures carry massive parameter sets. When sample periods are short, feature dimensions are high, or parameters are repeatedly tuned, the model easily picks up random noise and accidental regularities in historical data. Such models often perform exceptionally well in backtests but show poor generalization ability in future markets. As Bagnara pointed out, machine learning research in asset pricing and return forecasting must guard against false discoveries and a lack of interpretability brought by complex models [13], a caution that applies equally to LSTM-based stock prediction and portfolio research.

Second, data leakage results in grossly overstated outcomes. Common sources of leakage are backtesting historical periods with real-time lists of index constituents, not removing delisted stocks, using financial data that is not available at each observation timestamp, standardizing features using the full sample, and iteratively selecting the model based on information from the test set. If a model uses information that real investors could not have known at the time, its backtest performance will be overstated. Hence, applying point-in-time data is essential in financial machine learning studies.

Third, transaction costs are frequently underestimated. Long-short and market-neutral strategies usually involve frequent rebalancing, short selling, stock borrow fees, slippage, and market impact

DOI: 10.54254/2755-2721/2026.LD35447

costs. The study by Fischer and Krauss has already shown that transaction costs noticeably erode LSTM strategy performance [7]. Thus, reporting only pre-cost returns makes a strategy appear far more attractive than it actually is.

Fourth, there is a mismatch between predictive objectives and investment goals. Most LSTM architectures adopt MSE or MAE as loss functions for model training, yet portfolio construction prioritizes cross-sectional ranking power, Sharpe ratio, maximum drawdown and net post-transaction-cost returns. A model might have a low prediction error but weak ranking performance; conversely, it might predict prices accurately but fail to build a reliably profitable portfolio. Therefore, model training objectives should align as closely as possible with portfolio management goals.

Fifth, risk exposure control is insufficient. Although some papers construct long-short portfolios, they do not sufficiently control for market beta and industry and style factor exposures. If the long and short portfolios are not balanced on these dimensions, the strategy's returns may be driven by systematic risk rather than true alpha. In the case of market-neutral strategies, risk cancelation is not an optional appurtenance, but rather the foundational criterion by which to judge whether a strategy is truly working.

6. Future research directions

Future research should first improve model objective functions. Rather than training LSTMs solely on MSE or MAE, it would be better to adopt ranking loss, pairwise loss, Rank-IC, differentiable Sharpe ratios, or utility functions. Such loss targets are far more consistent with two core evaluation dimensions: cross-sectional ranking power and real portfolio profitability, which are the core pursuits of long-short and market-neutral trading strategies. A feasible experimental framework takes S&P 500 constituent equities as research samples, and trains four structurally identical LSTM models with distinct loss functions (MSE, ListNet ranking loss, pairwise loss and differentiable Sharpe ratio) under a 36-month rolling training window. At the end of each test period, construct a monthly rebalanced long-short portfolio sorted by predicted returns, and evaluate it using three metrics: Rank-IC, portfolio Sharpe ratio, and factor alpha. If the portfolio under the differentiable Sharpe ratio loss significantly outperforms the MSE baseline, it would demonstrate that improving the objective function can genuinely boost strategic value.

Secondly, LSTM with multi-source features should be considered in future work. Of course, OHLCV time series are readily available, but models based on these data alone lack in interpretability from an economic perspective. A more sensible strategy is to treat the price series, the technical indicators, the fundamental variables, the Fama-French factors, the momentum factors, the industry variables, the macroeconomic variables as a group and let the model find time-series patterns across each cross-sectional unit. The LSTM-based cross-section factor model proposed by Yao et al. In this direction, it focuses on employing LSTM to learn the cross-sectional factor structure of stock returns [14].

Third, future research requires more rigorous backtesting designs. Studies can attempt to use point-in-time data combined with rolling out-of-sample validation or walk-forward validation instead of simply splitting training and test sets randomly. For example, Researchers are able to use all non-ST A-share equities as research samples and compare three backtesting frameworks based on point-in-time financial data from Wind or CSMAR, including random sample splitting, rolling validation that does not incorporate point-in-time adjustment, and rolling validation that has full access to point-in-time data. By observing the Sharpe ratio, maximum drawdown, and return after costs of these different strategies over both normal and market crisis periods, scholars can

DOI: 10.54254/2755-2721/2026.LD35447

quantitatively evaluate attitude towards data leakage and oversimplified backtesting are to inflate strategy returns. This minimizes information leakage and is more comparable to a real-world investment environment, therefore. Meanwhile, research should also include crisis-period testing, cross-market testing, and robustness checks to determine if the model only works during specific market phases.

Fourth, future research should place greater emphasis on transaction costs and implementability. Studies should report both returns before and after costs, as well as sensitivity analyses for transaction costs. At the same time, practical concerns about liquidity, short-sale constraints, rates of turnover, and capacity constraints ought be raised. Jensen et al. research on the implementable efficient frontier suggests that machine learning can be used not only for return prediction but also for learning portfolio weights after considering transaction costs and execution constraints [15].

Fifth, future research should strengthen risk neutralization and performance attribution. Market neutrality goes far beyond balancing the total notional value of long and short positions; crucially, it imposes strict constraints on portfolio market beta, industry tilts and style factor exposures. Strategy results should also undergo attribution analysis to distinguish whether returns originate from market risk, traditional factors, or independent alpha signals generated by the LSTM.

7. Conclusion

In general, the optimal positioning of LSTM in stock portfolio construction is not to predict standalone stock prices, but to provide alpha signals for long-short and market-neutral strategies. Existing research shows that LSTM architectures are capable of capturing nonlinear correlations and long-range temporal dependencies embedded in financial time series, and their forecasting outputs can be integrated into investment workflows through equity ranking derived from forecasted returns and subsequent long-short portfolio building. Correspondingly, research has gradually evolved from early price prediction to cross-sectional ranking and market-neutral portfolio construction.

However, current studies cannot easily prove that LSTM can consistently beat the market. Overfitting, data leakage, underestimated transaction costs, insufficient control over risk exposure, and weak out-of-sample stability all compromise the reliability of conclusions. Therefore, the truly meaningful direction for LSTM-based stock prediction is to move away from pure price forecasting toward an integrated framework of "prediction - ranking - portfolio construction - risk control - trading execution - performance attribution". Under this framework, researchers need to implement point-in-time data, rolling out-of-sample validation, market-neutral constraints, transaction cost modeling, and factor attribution analysis to determine whether LSTM signals truly offer independent stock selection value. This paper also carries several inherent limitations: first, the reviewed literature is primarily drawn from English-language academic databases, which omits relevant research focusing on the Chinese market and other non-English speaking markets, where variations in institutional environments and investor structures could affect the efficacy of LSTM strategies; second, this paper does not conduct a cross-comparison of LSTM model hyperparameters (such as the number of hidden layer units, time window length, and learning rate) across different studies, despite their significant impact on model performance; Furthermore, this manuscript centers solely on LSTM applications under supervised learning paradigms, while neglecting the untapped potential of reinforcement learning and online learning algorithms for portfolio optimization. These emerging learning paradigms remain worthy of in-depth investigation in follow-up studies.

DOI: 10.54254/2755-2721/2026.LD35447

References

[1] Ozbayoglu, A. M., Gudelek, M. U., & Sezer, O. B. (2020). Deep learning for financial applications: A survey. *Applied Soft Computing, 93*, 106384.
[2] Jiang, W. (2021). Applications of deep learning in stock market prediction: Recent progress. *Expert Systems with Applications, 184*, 115537.
[3] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735-1780.
[4] Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56.
[5] Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance, 52*(1), 57-82.
[6] Jacobs, B. I., & Levy, K. N. (2005). *Market neutral strategies*. Wiley.
[7] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research, 270*(2), 654-669.
[8] Hou, X., Wang, K., Zhang, J., & Wei, Z. (2020). An enriched time-series forecasting framework for long-short portfolio strategy. *IEEE Access, 8*, 31992-32002.
[9] Cipiloglu Yildiz, Z., & Yildiz, S. B. (2022). A portfolio construction framework using LSTM-based stock markets forecasting. *International Journal of Finance & Economics, 27*(2), 2356-2366.
[10] Nafia, A., Yousfi, A., & Echaoui, A. (2023). Equity-market-neutral strategy portfolio construction using LSTM-based stock prediction and selection: An application to S&P500 consumer staples stocks. *International Journal of Financial Studies, 11*(2), 57.
[11] Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies, 33*(5), 2223–2273.
[12] Chen, L., Pelger, M., & Zhu, J. (2024). Deep learning in asset pricing. *Management Science, 70*(2), 714–750.
[13] Bagnara, M. (2024). Asset pricing and machine learning: A critical review. *Journal of Economic Surveys, 38*(1), 27–56.
[14] Yao, H., Xia, S., & Liu, H. (2024). Return predictability via a long short-term memory-based cross-section factor model: Evidence from Chinese stock market. *Journal of Forecasting, 43*(6), 1770–1794.
[15] Jensen, T. I., Kelly, B. T., Malamud, S., & Pedersen, L. H. (2024). Machine learning and the implementable efficient frontier. SSRN Working Paper.