Proceedings of CONF-SEML 2026 Symposium: Computational Analysis and Modeling in Complex Intelligent Syste (CONF-SEML 2026)
DOI: 10.54254/2755-2721/2026.GU34723

Review on the Integration of Multi-Factor Models and Artificial Intelligence for Quantitative Trading

Yang Zhang

Faculty of Management, Guangxi Minzu University, Nanning, China
13397852137@163.com

Abstract. Artificial intelligence has become a key tool in quantitative trading, while the multi-factor model remains a common framework for stock selection and asset allocation. Existing reviews, however, usually discuss multi-factor models or AI-based quantitative strategies separately. They rarely provide a systematic account of how the two approaches are integrated. This paper aims to fill this gap by reviewing recent Chinese and international studies and examining the integration of multi-factor models with artificial intelligence in quantitative trading. The review shows that traditional multi-factor models provide interpretable variables and economic logic, but they have limitations in factor construction, nonlinear modelling and market regime adaptation. Machine learning improves factor screening and return prediction; deep learning extends the use of time-series and textual information; reinforcement learning further links prediction results with portfolio adjustment and trading decisions. These developments increase the flexibility of quantitative strategies, but they also bring problems such as overfitting, weak interpretability, unstable generalization and regulatory risk. Future research should give more attention to explainable AI, dynamic factor selection, multi-source data fusion and risk-constrained trading systems. Such work can help connect economic meaning with data-driven modelling and improve the practical value of AI-enhanced multi-factor strategies.

Keywords: multi-factor model, artificial intelligence, quantitative trading, machine learning, risk control

1. Introduction

China's A-share market has become more institutionalized, but it has not fully reached a semi-strong efficient state. Public information, fundamental indicators and trading behavior may still create opportunities for excess returns [1]. In this context, quantitative trading has become an important way to convert market information into investment rules. The multi-factor model is one of its core frameworks. It explains stock returns through valuation, growth, profitability, turnover, momentum, sentiment and other variables. For example, Lin Zhangyi used CSI 300 constituents as the sample and found that an equal-weighted Z-score multi-factor model performed better than the market benchmark [2].

However, traditional multi-factor models are no longer sufficient on their own. Their factors are often selected by researchers in advance, and many models rely on relatively stable linear relations.

DOI: 10.54254/2755-2721/2026.GU34723

In real markets, factor effectiveness changes with financial cycles, sector rotation and investor sentiment. Liu Yuxuan, Jin Weize and Yuan Liang show that adding financial cycle indicators can improve the performance of multi-factor stock selection [3]. Xia Zhendong and co-authors also find that adding a sentiment factor to the Fama-French three-factor model increases its explanatory power [4]. These findings suggest that factor models need to respond to changing market states and richer data sources.

This necessity underscores the integration of multi-factor models and artificial intelligence. Multi-factor models provide economic meaning and a clear structure for strategy design, but they are weak in handling high-dimensional variables, nonlinear relations and unstructured data. Artificial intelligence can partly fill this gap. Machine learning helps screen factors and capture complex interactions. Deep learning can process financial time series and textual information. Reinforcement learning can connect prediction with trading actions and risk-return adjustment. Recent research on multifactor timing shows that deep learning is more useful when it is constrained by economic structure. Cotturo, Liu and Proner combine multitask learning with LSTM to capture common factor structure and macro-financial states, which directly supports the need to integrate AI with factor models [5]. Algorithmic trading research also shows that data-driven trading can affect market pricing efficiency through volatility and investor belief channels [6]. The key issue is not whether AI should replace factor models, but how AI can strengthen factor selection, prediction and decision-making while retaining the economic logic of factors.

The research gap is explicit. Existing studies and reviews usually examine multi-factor models, investor sentiment, machine learning, deep learning or algorithmic trading as separate topics. They lack sufficient systematic discussion on the integration of multi-factor models and AI methods in quantitative trading. The links among factor construction, AI-based prediction, trading execution, risk control and financial regulation remain fragmented [7]. This makes it difficult to understand the full application path of the integrated approach.

This paper aims to clarify how multi-factor models and artificial intelligence can be integrated in quantitative trading. It focuses on three questions: how AI improves factor screening and return prediction, how textual and market-state information can be included in factor systems, and how intelligent models support strategy optimization and risk control. The value of this study is to connect the economic logic of factors with data-driven modelling. It can provide a clearer reference for explainable quantitative strategies, dynamic factor selection and risk-aware intelligent trading systems.

2. Theoretical background

The basic idea of the multi-factor model is to explain stock return differences through several risk factors or characteristic variables. It does not attribute stock price changes to a single indicator. Instead, it considers valuation, profitability, growth, size, turnover, momentum and industry information at the same time. Compared with a single-factor model, a multi-factor model can screen stocks from several dimensions and reduce the bias caused by the short-term failure of one factor. Lin Zhangyi's equal-weighted Z-score model based on CSI 300 constituents showed that a multi-factor portfolio could outperform the market benchmark to some extent [2]. This finding supports the use of multi-factor models as a clear rule-based foundation for quantitative stock selection.

In asset pricing research, multi-factor models also explain the sources of risk. Three-factor, five-factor and six-factor models try to include market return, size, value, profitability, investment style and other variables in a unified framework. Zhan Kai and Ren Yawei tested the Barillas-Shanken six-factor model in China's stock market. They found that an expanded factor model could improve

DOI: 10.54254/2755-2721/2026.GU34723

the explanation of cross-sectional return differences [8]. Factor effectiveness is not fixed. Liu Yuxuan and co-authors introduced a financial cycle indicator into a factor model and showed that macro-financial conditions affect model return and stability [3]. This means that factor models need to respond to market regimes rather than rely on a constant factor structure.

Artificial intelligence provides tools for dynamic adjustment. Machine learning can select stable signals from high-dimensional variables. Deep learning can extract complex features from price series, announcement texts and sentiment data. Gu, Kelly and Xiu show that machine learning can improve empirical asset pricing by processing high-dimensional predictors and capturing nonlinear interactions that traditional regression models often miss [9]. Cotturo, Liu and Proner show that deep learning can improve multifactor timing when it combines common factor structure with LSTM-based macro-financial state extraction [5].

In this paper, integration does not mean that artificial intelligence completely replaces traditional factor models. It means placing the economic meaning of factors, market state variables and data-driven algorithms in the same analytical framework. Multi-factor models provide the explanatory base, while artificial intelligence provides feature learning and prediction tools. Research on investor sentiment also shows that non-rational variables outside traditional financial factors affect asset pricing. Xia Zhendong and co-authors added a sentiment factor to the Fama-French three-factor model and improved its explanatory power [4]. Xiong Yuning constructed multi-channel investor sentiment from trading markets, online news and social media, and introduced it into a quantitative stock selection strategy [10]. These findings provide the theoretical basis for the following literature review.

3. Literature review

3.1. Traditional multi-factor models

Traditional multi-factor models remain the starting point of quantitative stock selection because they connect returns with interpretable factors. Their main advantage is not only prediction, but also explanation. Valuation, growth, profitability, turnover and other factors allow investors to understand where returns may come from. Studies based on CSI 300 and A-share samples show that factor screening, redundancy testing and backtesting have formed a mature basic process [2]. Similar work also confirms that traditional factor strategies can be implemented in Chinese market conditions [11]. The following discussion therefore examines three core weaknesses of traditional multi-factor models in more detail.

The first limitation is the reliance on manually designed factors and relatively stable linear relations. Extended asset-pricing models improve explanatory power, but they still depend on predefined factor structures [8]. Sentiment factors and Black-Litterman allocation models show that behavioural information can be added to this framework, yet the model still requires researchers to decide what information should enter the factor system [4]. This makes the framework clear and interpretable, but less flexible when new data types or hidden interactions appear.

The second limitation is instability across time and market states. Multi-factor models can support dynamic allocation beyond stock selection [12]. They are also useful in practical quantitative investment [13]. However, industry rotation, financial cycles and sector differences show that factor weights and factor effectiveness may change when the macro-financial environment changes [3]. Wan Xiangyu and Zhang Chen further indicate that factor loadings vary across sectors and periods [14]. Fixed weights may therefore weaken model stability.

DOI: 10.54254/2755-2721/2026.GU34723

The third limitation is the narrow information boundary. Traditional models are strong when financial statements and price-volume data are clean, but they are weaker in capturing investor attention, sentiment shocks and nonlinear behavioural responses. Industry-rotation research improves this problem by linking factor models with economic stages [15]. Sentiment-based allocation also shows the value of behavioural variables [16]. Overall, traditional multi-factor models offer an interpretable foundation, but their manual construction, linear logic and limited data scope create the need for AI-based optimization.

3.2. Machine learning optimization

Machine learning optimizes multi-factor models mainly through factor screening, nonlinear modelling and model combination. Compared with scoring methods or linear regression, it is stronger at processing high-dimensional factors and interaction effects, but it does not automatically improve strategy quality. Its effect depends on the match between the algorithm, the factor structure and the market scenario. Gu, Kelly and Xiu show that machine learning improves empirical asset pricing mainly because it can process high-dimensional predictors and capture nonlinear interactions that traditional regression models often miss [9]. This means that algorithm choice should depend on the structure of the factor problem, rather than on a general belief that more complex models are always better.

Support vector machines are more suitable for medium-sized samples and classification tasks where the boundary between high-return and low-return stocks is nonlinear. Their advantage is stability under a clear margin, yet they are sensitive to the kernel function, parameter setting and feature scaling [17]. Logistic regression is easier to explain, but it is often too simple for complex factor structures. Random forest is more robust to noise than a single tree, but its averaging mechanism may weaken precise timing signals [18]. The combined use of SVM timing and XGBoost stock selection suggests a more practical route: different algorithms can be assigned to different stages rather than treated as one universal model [19].

Boosting algorithms are better suited to large factor pools with nonlinear interactions. LightGBM is efficient in high-dimensional settings and has lower training cost [20]. XGBoost is strong in prediction accuracy and can capture complex split relations among factors [21]. Compared with SVM, boosting models are better for large-scale feature engineering; compared with linear models, they improve predictive power. Their weakness is weaker economic interpretability and a higher risk of fitting noise when validation is loose.

Stacking and neural networks move one step further. Stacking improves robustness by combining base models, but it increases model complexity and makes error sources harder to trace [22]. Neural networks can learn high-order nonlinear relations, yet they need larger samples and are less transparent [23]. Elastic Net is different: it is less powerful in nonlinear fitting, but it is more interpretable and useful for reducing factor redundancy [24]. Boosting methods can strengthen factor combinations, but they also depend heavily on parameter tuning and data quality [25]. Overall, machine learning should be understood as a set of scenario-specific tools rather than a single replacement for traditional multi-factor models.

3.3. Deep learning, sentiment factors and reinforcement learning strategies

Deep learning, sentiment factors and reinforcement learning should be understood as three different extensions of quantitative trading. Deep learning mainly deals with time-varying market states and complex nonlinear patterns. Cotturo, Liu and Proner show that a dynamic multitask neural network

DOI: 10.54254/2755-2721/2026.GU34723

can combine common factor structure with LSTM-based macro-financial state extraction, which makes deep learning more suitable for multifactor timing than ordinary static models [5]. In time-series applications, deep learning is not limited to stock price prediction. Zhang, Zhang, Cucuringu and Qian show that neural networks can improve intraday realized volatility forecasting by exploiting volatility commonality across stocks and market volatility information [26]. The weakness is clear: these models are data-hungry, less interpretable and sensitive to training design.

Sentiment factors fill a behavioural-finance gap left by traditional factor models. Financial statements and price-volume indicators cannot fully explain short-term attention, media pressure and investor emotion. Multi-channel sentiment research shows that information from trading markets, online news and social media can enrich the factor system [10]. Text analysis based on Weibo comments further links public opinion factors with Attention-LSTM and trading strategies [27]. The advantage of sentiment factors is that they capture behavioural information earlier than some financial indicators. Their risk is noise, because online attention does not always represent stable returns information.

Reinforcement learning moves the research focus from prediction to decision. It does not only ask which stock may rise, but also learns when to buy, sell, hold or rebalance. SAC-based research uses improved experience replay and risk indicators to increase adaptability under changing market styles [28]. DQN-based risk-control research adds Sharpe ratio and maximum drawdown into the reward function, which makes trading decisions more risk-aware [29]. Sarsa-based pairs trading replaces fixed parameters with dynamic optimization, showing how reinforcement learning can support strategy execution [30].

At the portfolio level, reinforcement learning also supports asset allocation and rebalancing. Strategies based on DDPG, A2C and PPO show that intelligent agents can adapt to different market conditions and improve risk-return performance [31]. Compared with traditional multi-factor and machine learning models, reinforcement learning is closer to real trading decisions. Its main limits are unstable training, reward-function design, transaction costs and regulatory constraints. This means that AI integration should not only pursue higher prediction accuracy. It should also connect factor meaning, model transparency and risk control.

4. Discussion

Several ways exist to connect multi-factor models and artificial intelligence, as demonstrated by the literature review. The more important issue is whether this connection improves trading logic, rather than only raising backtest performance. The first limitation is the tension between economic meaning and data-driven fitting. Traditional factors are useful because they give a clear interpretation of return sources. AI models can combine factors more flexibly, but this flexibility may also weaken the economic meaning of the strategy. If a model gives high weight to many unstable signals, it may become difficult to tell whether the return comes from value, growth, momentum, sentiment or sample noise. Since factor loadings change across sectors and periods, an integrated strategy should not treat AI as a tool for automatically adding more variables [16]. It should use AI to test which factors remain robust under changing market states.

The second challenge is hidden behind the apparent advantage of AI algorithms. Machine learning can capture nonlinear relations and factor interactions, but this does not mean that more complex models are always better than simpler ones. Elastic Net has value when the factor pool is large because it can reduce redundant variables and keep the model relatively interpretable [24]. Tree-based models such as XGBoost and LightGBM are better suited to high-dimensional tabular factors, yet their performance can depend heavily on sample division and parameter tuning. Neural

DOI: 10.54254/2755-2721/2026.GU34723

networks and deep learning can process time-varying market states and complex nonlinear patterns, while sentiment studies show that news and social media data can enrich the factor system [5, 10]. For this reason, model comparison should not rely only on accuracy, excess return or Sharpe ratio. It should also examine stability across market regimes, factor contribution, turnover and the cost of wrong signals.

A third issue is the gap between research design and actual trading. Many empirical studies use clean historical data and assume smooth execution. Real trading is less ideal. Recent work on intraday volatility forecasting also shows that machine learning can support risk assessment and pre-trade cost analysis in automated trading, which makes practical execution an important part of AI-based quantitative strategies [26]. Financial statements may be delayed, sentiment data may contain noise, and high-turnover strategies face transaction costs, slippage and liquidity limits. These frictions can reduce or even remove the excess return shown in backtests. Reinforcement learning makes this problem more visible. It can learn buy, sell, hold and rebalance actions, but the result depends strongly on the reward function. If the reward only focuses on return, the model may overtrade or take excessive risk. The studies that include maximum drawdown, Sharpe ratio and other risk indicators show a more realistic direction for intelligent trading systems [28, 29].

These problems suggest several practical rules for future integration. First, AI should be used to improve factor selection, not to replace all economic reasoning. Each strategy still needs a clear explanation of why the selected factors should work. Second, backtesting should include walk-forward validation, transaction-cost adjustment, liquidity filters and stress tests across bull, bear and volatile markets. Third, alternative data such as textual sentiment should be treated as supplementary information, not as a complete substitute for price and fundamental data. Fourth, algorithmic execution and regulatory constraints should be included in strategy design. Algorithmic trading may improve pricing efficiency, but it also raises concerns about market stability and monitoring [6]. Big-data trading further requires reliable data governance, privacy protection and regulatory coordination [7]. In this sense, the real value of integrating multi-factor models with AI lies in building explainable, risk-aware and executable strategies, rather than simply producing more complex models.

5. Conclusion

This paper reviews the integration of multi-factor models and artificial intelligence in quantitative trading. Overall, the multi-factor model remains an important framework for explaining asset returns, building stock selection strategies and managing portfolios. Its strengths lie in clear logic, interpretable indicators, convenient backtesting and risk control. As the A-share market changes and data dimensions increase, traditional models also show several limitations. These include strong linear assumptions, unstable factor effectiveness and weak ability to process unstructured data.

Artificial intelligence expands the technical boundary of this framework. Machine learning, deep learning and reinforcement learning improve traditional models from the levels of factor screening, feature representation and trading decisions. Quantitative trading is therefore moving from fixed rules toward dynamic learning. When the two are combined, the model can preserve the explanatory logic of factors while improving its response to complex data and market change.

However, integration does not mean that a more complex model is always better. Artificial intelligence models face overfitting, weak interpretability, limited transferability and transaction cost constraints. Quantitative strategies are also affected by market style, regulatory environment and investor behavior. Future research should improve prediction ability while paying more attention to model transparency, risk control and practical execution. A more robust path is to combine the

DOI: 10.54254/2755-2721/2026.GU34723

interpretability of multi-factor models with the learning ability of artificial intelligence, especially in dynamic factor selection, multi-source data fusion and risk-constrained trading strategies.

References

[1] Guo, Q. (2021). Research of Quantification Based on Technical and Fundamental Factors and Test on Efficient Market Theory [Master's thesis, China University of Petroleum (Beijing)].
[2] Lin, Z.Y. (2020). Quantitative Stock Picking Strategy Based on Multi-Factor Model: A Case Study of CSI 300 Index [Master's thesis, Guangdong University of Finance and Economics].
[3] Liu, Y.X., Jin, W.Z. and Yuan, L. (2022). Optimization and empirical research of a multi-factor quantitative stock selection model: Introducing financial cycle indicators. Price: Theory & Practice, 2022(4), 141-145.
[4] Xia, Z.D., Li, W., Wang, H. and Lin, X.K. (2021). Risk factor pricing based on investor sentiment. Investment and Entrepreneurship, 32(3), 4-6.
[5] Cotturo, P., Liu, F., & Proner, R. (2026). Multifactor timing with deep learning. Journal of Financial Econometrics, 24(3), nbag006. https: //doi.org/10.1093/jjfinec/nbag006.
[6] Wang, F. (2022). Does algorithmic trading improve the pricing efficiency of China's stock market?. South China Finance, 2022(8), 49-64.
[7] Luo, S.J., Gu, L.M., Lin, L.K. and Li, Q.Y. (2022). Quantitative trading strategies and financial regulation based on big data technology. Fortune Today, 2022(5), 13-15.
[8] Zhan, K. and Ren, Y.W. (2021). Applicability of the Barillas-Shanken six-factor model in China's stock market. Journal of Shaoguan University (Social Science), 42(4), 60-66.
[9] Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. The Review of Financial Studies, 33(5), 2223–2273. https: //doi.org/10.1093/rfs/hhaa009.
[10] Xiong, Y.N. (2021). Research on Multi-Channel Investor Sentiment Analysis Methods for the Stock Market [Master's thesis, Southwestern University of Finance and Economics].
[11] Zou, J.Q. (2020). Construction and Research of a Multi-Factor Quantitative Model in the A-Share Market [Master's thesis, Shanghai University of Finance and Economics].
[12] Xu, M.P. and Liang, R. (2021). Dynamic asset allocation based on an extended multi-factor model: Evidence from Shanghai A-shares. Mathematics in Practice and Theory, 51(4), 35-43.
[13] Zou, L.X. (2021). Research on quantitative investment based on a multi-factor model. China Journal of Commerce, 2021(24), 100-103.
[14] Wan, X.Y. and Zhang, C. (2022). Time-varying characteristics of factor pricing models and sector differences in the stock market: Estimation based on a time-varying parameter seemingly unrelated method. Statistics & Decision, 2022(4), 154-158.
[15] Luo, J. (2022). Empirical research on multi-factor stock selection model and investment effect based on industry rotation strategy. Science & Technology Information, 20(20), 148-152.
[16] Pang, J. (2021). Industry asset allocation under the Black-Litterman model: Combining investor sentiment index. Science Research Management, 42(6), 8-24.
[17] Yin, W.C. and Chu, Q.Z. (2021). Multi-factor stock selection modelling and application based on support vector machine. Mathematical Modeling and Its Applications, 10(4), 64-71.
[18] Liu, Y. (2021). Research on Quantitative Stock Selection Models Based on Machine Learning [Master's thesis, Xidian University].
[19] Guan, R.J. (2020). Research on Quantitative Investment Stock Selection Based on Machine Learning [Master's thesis, Southwestern University of Finance and Economics].
[20] Wu, B.X. (2021). Research on Multi-Factor Quantitative Stock Selection Strategy Based on LightGBM [Master's thesis, Northwest University].
[21] Zhu, Y.B. (2020). Design of a Multi-Factor Stock Selection Scheme Based on XGBoost and LightGBM Algorithms [Master's thesis, Nanjing University].
[22] Luo, Z.N. (2021). Research on a machine learning multi-factor stock selection model based on Stacking ensemble. China Price, 2021(11), 77-79.
[23] Liu, M.Y. and Pang, H.L. (2022). Neural network multi-factor stock selection model. Journal of Changchun University of Technology, 43(2), 152-158.
[24] Shu, S.K. and Li, L. (2021). Multi-factor stock selection strategy based on Elastic Net penalty. Statistics & Decision, 2021(16), 157-161.

DOI: 10.54254/2755-2721/2026.GU34723

[25] Wang, Y.Z. (2020). Empirical Research on Multi-Factor Quantitative Stock Selection Based on Boosting Algorithms [Master's thesis, Shandong University].
[26] Zhang, C., Zhang, Y., Cucuringu, M., & Qian, Z. (2024). Volatility forecasting with machine learning and intraday commonality. Journal of Financial Econometrics, 22(2), 492–530. https: //doi.org/10.1093/jjfinec/nbad005.
[27] Wang, P. (2021). Design and Implementation of a Stock Trading Strategy Based on Text Analysis and Reinforcement Learning Technology [Master's thesis, Southwestern University of Finance and Economics].
[28] Liu, P.Y. (2021). Improved Reinforcement Learning Algorithm and Its Application in Quantitative Trading [Master's thesis, Northeast Petroleum University].
[29] Wang, W.P. (2021). Risk control algorithm for a trading model based on deep reinforcement learning. Modern Computer, 2021(3), 42-47.
[30] Li, J. (2021). Design of a Commodity Futures Pairs Trading Strategy Based on Reinforcement Learning Algorithms [Master's thesis, Shanghai Normal University].
[31] Fan, X.Y. (2021). Research on Portfolio Strategy Based on Deep Reinforcement Learning [Master's thesis, Dalian University of Technology].