Eurasia Journal of Science and Technology
ISSN: 2663-1016
DOI: https://doi.org/10.61784/ejst3070

A-SHARE INTELLIGENT STOCK SELECTION STRATEGY BASED ON THE DEEPSEEK LARGE MODEL: TECHNICAL ROUTES, FACTOR SYSTEMS, AND EMPIRICAL RESEARCH

HaiLong Liao
School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China.
Corresponding Email: jnhailong@126.com

Abstract: This research focuses on the application of the DeepSeek large - scale model in intelligent stock selection in the A - share market. It constructs a multi - dimensional factor analysis framework that integrates reinforcement learning and a mixture - of - experts architecture. Through empirical research, the advantages of this model in terms of return acquisition and risk control are verified, providing new intelligent strategies for A - share investment and promoting technological innovation and development in the capital market.
Keywords: DeepSeek large - scale model; Intelligent stock selection in the A - share market; Factor system; Empirical research

1 INTRODUCTION

The global financial market is in a period of rapid transformation, with increasingly prominent complexity and information explosion characteristics. Against this background, traditional quantitative models face numerous challenges in investment decision - making. For example, the Fama - French three - factor model [1] has an explanatory power of only 37.2% for the TMT sector, highlighting the insufficiency of traditional quantitative models in capturing non - linear correlations. According to a report by a foreign research institution cited in The New Yorker magazine, ChatGPT responds to approximately 200 million requests per day and consumes over 500,000 kWh of electricity in the process. That is to say, ChatGPT's daily electricity consumption is equivalent to that of 17,000 American households [2]. The high processing cost limits its application in high - frequency trading scenarios. In addition, the market changes dynamically and frequently. In 2023, the industry rotation frequency of the A - share market increased by 28% year - on - year. Traditional models lack an effective dynamic environment adaptation mechanism and are difficult to keep up with the pace of market changes, greatly reducing the timeliness and accuracy of investment decisions.

The DeepSeek technology system, with its architectural innovation (MLA multi - head latent attention mechanism) and breakthrough in training paradigms (GRPO reinforcement learning algorithm), brings new opportunities to solve the above - mentioned problems [3]. It shows unique advantages in computational efficiency, resource allocation, and decision - making path optimization, injecting new vitality into the financial investment field.

This research focuses on the application of the DeepSeek large - scale model[4] in intelligent stock selection in the A - share market. It constructs a multi - dimensional factor analysis framework that integrates reinforcement learning and a mixture - of - experts architecture. This research has multiple important values. On the one hand, it is the first to systematically verify the implementation effectiveness of Chinese open - source large models in complex financial scenarios, filling a research gap in the relevant field. By constructing an interpretable AI - enhanced factor system, it provides investors with a more scientific and comprehensive investment analysis perspective. On the other hand, it demonstrates the optimization effect of edge - side deployment on the investment decision - making chain. Experiments show that the inference latency of the 7B model is less than 50 ms, which is of great significance for improving the efficiency of investment decision - making. The results of this research provide empirical support for the popularization of AI technology, and are expected to promote the transformation of the Chinese capital market towards intelligence, helping small and medium - sized investors gain a more favorable investment position in a complex market environment.

2 HISTORICAL DATA AND TREND ANALYSIS OF THE A - SHARE MARKET

2.1 Review of Historical Data

Since its establishment, the A - share market has gone through multiple development stages. The market scale has been continuously expanding, the investor structure has been gradually optimized, and the market system has become increasingly perfect. Reviewing historical data can provide rich reference for the current research.

Looking at the overall market trend, the Shanghai Composite Index has shown obvious cyclical fluctuations over the past few decades. For example, during the period from 2005 to 2007, the Shanghai Composite Index soared from 998 points to 6124 points, with a significant bull market. This was mainly due to factors such as the institutional dividend released by the share - splitting reform, the high - speed growth of the macro - economy, and a large amount of capital flowing into the stock market. Subsequently, in 2008, affected by the global financial crisis, the Shanghai Composite

Index dropped sharply to 1664 points. Market panic spread, and investors' confidence was frustrated. In 2014 - 2015, there was another rapid upward trend. The Shanghai Composite Index rose from around 2000 points to 5178 points. This stage was closely related to factors such as loose monetary policy and the development of margin trading business. However, the market then adjusted rapidly, resulting in a stock market crash and causing huge losses to investors.
In terms of industry performance, the price changes of different industries vary significantly in different periods. In the early stage, traditional energy and financial industries dominated the market. With the adjustment of the economic structure and the development of technology, consumer and technology industries have gradually emerged. For example, during 2019 - 2020, the liquor industry benefited from consumption upgrading and increased brand concentration. The stock prices of leading enterprises such as Kweichow Moutai and Wuliangye increased significantly. The semiconductor industry became a market hotspot against the background of national policy support and strong domestic substitution demand. Enterprises such as SMIC and GigaDevice Semiconductor performed outstandingly.

2.2 Trend Analysis

In recent years, the A - share market has shown some new trends. First, the degree of market institutionalization has been continuously increasing. With the continuous inflow of long - term funds such as social security funds, pension funds, and foreign capital, as well as the continuous expansion of the scale of public and private funds, institutional investors have gradually strengthened their say in the market. This makes the market investment style pay more attention to fundamental analysis and long - term investment value, and the stock price trend becomes more rational.
Second, scientific and technological innovation has become the core driving force of the market. Under the background of the country's strong promotion of the innovation - driven development strategy, the technology industry has developed rapidly, and its weight in the A - share market has been continuously increasing. From 5G communication, artificial intelligence to new energy vehicles, related industrial chain enterprises have emerged in an endless stream, bringing new investment opportunities to the market. For example, the new energy vehicle industry, driven by multiple factors such as policy subsidies, technological progress, and growing market demand, has seen a significant increase in the market value of enterprises such as BYD and Contemporary Amperex Technology Co., Limited, making the entire industry sector the focus of the market.
Third, the correlation between the market and the macro - economy has become closer. Changes in macro - economic data, such as GDP growth rate, inflation rate, and monetary policy adjustments, have a more significant impact on the A - share market. In the economic recovery stage, the market often anticipates an increase in corporate earnings, and stock prices rise. When there is greater downward pressure on the economy, the market is more cautious, and stock prices fluctuate more violently.

3 DEEPSEEK TECHNICAL FOUNDATION AND FINANCIAL SCENARIO ADAPTABILITY

3.1 Analysis of Core Technologies

The technical advantages of DeepSeek are reflected in multiple dimensions, mainly including three aspects: computational efficiency innovation, dynamic resource allocation [5], and decision - making path optimization [6].
In terms of computational efficiency innovation, the MLA architecture is the key. When processing large - scale data, traditional Transformer models have high KV cache requirements, which limit their processing efficiency to a certain extent. The MLA architecture effectively reduces the KV cache requirements through implicit attention mapping. When processing input sequences of different lengths, the MLA technology can effectively reduce the cache size, thereby improving the performance of the model. For example, in DeepSeek - V3, after using the MLA technology, the compression ratio of the KV cache reached 6 times. This means that when processing text sequences of the same length, the MLA technology can significantly reduce the required memory space, thus improving the operation efficiency of the model [7]. This improvement enables DeepSeek to support the real - time processing of high - frequency market data. In the rapidly changing financial market, it can timely capture market information and provide timely data support for investment decisions.
The DeepSeekMoE mixture - of - experts model realizes dynamic resource allocation. The model adopts a dynamic parameter activation mechanism. DeepSeekMoE is an innovative mixture - of - experts (MoE) architecture, aiming to achieve higher expert specialization and computational efficiency through fine - grained expert segmentation and shared expert isolation strategies. DeepSeekMoE further divides experts, enabling each expert to focus more on specific knowledge fields or tasks. This fine - grained division allows the model to improve the effect and efficiency by flexibly combining multiple experts when dealing with complex tasks. For example, in the DeepSeekMoE 16B model, 8 experts are selected from 64 experts for activation, thereby achieving higher knowledge acquisition accuracy and computational efficiency [8]. This mechanism, while ensuring the model capacity, significantly reduces the inference energy consumption by 70%. In financial scenarios, a large number of computational tasks require huge computational resources. This feature of the DeepSeekMoE model enables more efficient completion of complex financial analysis tasks under limited computational resource conditions.
Decision - making path optimization is another important advantage of DeepSeek. The GRPO (Generalized Reward Policy Optimization) algorithm allows the model to autonomously optimize investment strategies in an unsupervised environment by designing a multi - objective reward function [9]. In the backtest of the R1 version, its self - correction accuracy rate is as high as 82.3%. This means that the model can continuously adjust investment strategies according to

market changes, improving the accuracy and adaptability of investment decisions. For example, when the market fluctuates, the model can adjust the position portfolio in a timely manner according to the feedback of the reward function, reducing risks and increasing returns.

3.2 Adaptability to Financial Scenarios

In the application of the A - share market, DeepSeek shows three characteristics: long - text parsing, real - time response ability, and a localized knowledge base.

Long - text parsing ability is crucial for financial investment. In the financial field, annual reports and prospectuses contain a large amount of key information. Accurately extracting this information is of great significance for investment analysis. DeepSeek performs well in this regard. Its accuracy rate for extracting key information from annual reports/prospectuses has increased to 89.7%, which is significantly higher than 86.2% of GPT - 4, and the processing speed has increased by 3.2 times. This enables investors to obtain the core information of enterprises more quickly and accurately, providing a strong basis for investment decisions.

Real - time response ability is another outstanding advantage of DeepSeek in financial scenarios. The distilled 7B model can perform 1200 factor calculations per second on a device with 12GB of video memory. This calculation speed can meet the strict real - time requirements of intraday trading. In intraday trading, the market changes rapidly. Timely factor calculation and investment decisions can seize fleeting investment opportunities and increase investment returns.

DeepSeek has also constructed a localized knowledge base. In response to the unique policy orientation of the A - share market, such as the screening of "specialized, refined, characteristic, and new" enterprises, a special fine - tuning data set has been constructed. This has increased the strategy specificity by 41.6%, enabling it to better adapt to the characteristics of the Chinese capital market and providing strong support for investors to explore high - quality investment targets that meet the policy orientation.

4 METHODOLOGY FOR CONSTRUCTING INTELLIGENT STOCK SELECTION MODELS

4.1 Data Layer: Multimodal Factor System

This research constructs a dynamic database containing 6 categories and 32 factors. These factors cover multiple dimensions such as value, growth, momentum, sentiment, industry chain, and policy, comprehensively reflecting the comprehensive situation of enterprises and the market environment.

Table 1 Multidimensional Indicators and AI Enhancement Methods for Each Factor Type

Factor Type	Typical Indicators	AI Enhancement Method
Value Factor	PE/PB/Dividend Yield	Industry - relative Valuation Deviation Correction
Growth Factor	ROE Growth Rate/R&D Expense Ratio	Patent Text Technical Barrier Assessment
Momentum Factor	20 - Day Volatility/Turnover Rate	Main Capital Flow Map Analysis
Sentiment Factor	Stock Forum Public Opinion Sentiment Value	R1 Model Inference Chain Confidence Weighting
Industry Chain Factor	Difference in Inventory Turnover Rates of Upstream and Downstream	V3 Code Parsing of Industry Database
Policy Factor	Degree of Matching with the "14th Five - Year Plan"	Semantic Embedding of the Government Work Report

In terms of value factors, typical indicators such as PE/PB and dividend yield are selected, and AI enhancement is carried out through industry - relative valuation deviation correction. When evaluating the value of an enterprise, although traditional PE/PB indicators are commonly used, they are easily affected by the overall industry valuation level. Through industry - relative valuation deviation correction, it is possible to more accurately judge the value position of an enterprise within the industry and avoid valuation misjudgments caused by industry - wide systematic factors.

Growth factors include indicators such as ROE growth rate and R&D expense ratio, and AI enhancement is carried out using patent text technical barrier assessment. R&D investment is an important driving force for the future development of an enterprise, and patents are an important manifestation of an enterprise's technical strength. By analyzing patent texts and evaluating the technical barriers of enterprises, it is possible to gain a deeper understanding of the growth potential of enterprises and provide a more comprehensive perspective for evaluating growth factors.

Typical indicators of momentum factors include 20 - day volatility and turnover rate, and AI enhancement is carried out with the help of main capital flow map analysis. The momentum effect of the market is of great significance in investment decision - making. The flow direction of main capital often indicates the short - term trend of the market. Through main capital flow map analysis, it is possible to more accurately grasp the market momentum and adjust investment strategies in a timely manner.

The sentiment factor takes the stock forum public opinion sentiment value as an indicator and performs AI enhancement through R1 model inference chain confidence weighting. The public opinion in stock forums reflects the emotions and expectations of market participants. Although it contains a large amount of noise information, through the processing of the R1 model and confidence weighting, valuable sentiment signals can be extracted to provide a reference for investment decisions.

The industry chain factor selects the difference in inventory turnover rates of upstream and downstream as an indicator and uses V3 code parsing of the industry database for AI enhancement. In the industry chain, changes in the inventory turnover rates of upstream and downstream enterprises reflect the supply - demand relationship and operation efficiency of the industry chain. Through V3 code parsing of the industry database, it is possible to conduct a more in - depth analysis of industry chain factors and explore investment opportunities on the industry chain.

The policy factor takes the degree of matching with the "14th Five - Year Plan" as an indicator and performs AI enhancement through semantic embedding of the government work report. Policies have a profound impact on the capital market. The "14th Five - Year Plan" clarifies the country's development strategy and key support areas. Through semantic embedding of the government work report, it is possible to more accurately evaluate the fit between enterprises and policies and discover policy - driven investment opportunities.

4.2 Model Layer: Three - Stage Training Framework

The model layer adopts a three - stage training framework, including supervised fine - tuning (SFT), reinforcement learning with human feedback (RLHF), and adversarial training. Each stage plays a key role in improving the performance of the model.

In the supervised fine - tuning (SFT) stage, historical data of the TOP50 portfolio from 2018 to 2022 (including 30 basic factors) is input, and a preliminary stock selection probability distribution is output. At this time, the Hit@10 accuracy rate reaches 68.4%. In this stage, historical data is used to train the model, allowing the model to learn basic investment rules and data characteristics, laying a foundation for subsequent training.

In the reinforcement learning with human feedback (RLHF) stage, a reasonable reward function is designed: Sharpe ratio - maximum drawdown - industry diversification. The Sharpe ratio reflects the risk - adjusted return of an investment portfolio, the maximum drawdown reflects the maximum loss that an investment portfolio may face, and the industry diversification measures the distribution of the investment portfolio in different industries to avoid risks caused by over - concentrated investment. The Monte Carlo tree search is used to explore the position portfolio space, and the GRPO algorithm plays an important role in this process, reducing the annualized return volatility by 22%. In this stage, the model continuously tries different investment strategies in a simulated market environment and optimizes the strategies according to the feedback of the reward function, improving the model's investment decision - making ability.

In the adversarial training stage, in order to improve the robustness of the model, 10% noise data (such as financial fraud features, abnormal trading volume patterns) is injected into the data, and a discriminator network is constructed. Through adversarial training, the overfitting rate of the model is reduced from 17.3% to 5.1%, effectively enhancing the stability and reliability of the model when facing complex market environments and abnormal data.

4.3 Deployment Layer: Cloud - Edge - End Collaborative Architecture

The deployment layer adopts a cloud - edge - end collaborative architecture, giving full play to the advantages of different levels to achieve efficient investment decision - making support.

In terms of cloud - based training, the 70B model updates the industry knowledge base monthly, and this process consumes approximately $1200 in computing power costs. The cloud has powerful computing resources and can support large - scale model training and data processing. By updating the industry knowledge base monthly, the model can timely obtain the latest industry information and market dynamics, maintaining sensitivity to market changes.

Edge inference deploys the 7B model at the brokerage trading terminal, with a latency of less than 100 ms. The trading terminal has extremely high requirements for real - time performance. The low - latency feature of edge inference enables investors to obtain the analysis results of the model in a timely manner during the trading process and make quick decisions.

In order to better adapt to market changes, a dynamic weight mechanism is also adopted, which automatically adjusts the factor weights according to market volatility. For example, during high - volatility periods, the weight of the momentum factor is increased by 15%. Changes in market volatility reflect the degree of market risk and uncertainty. By dynamically adjusting factor weights, it is possible to make more rational use of various factors in different market environments, improving the adaptability and effectiveness of investment strategies.

5 Empirical Analysis: Strategy Backtesting and Market Impact

5.1 Backtesting Results (January 2023 - June 2024)

Backtesting was conducted on the DeepSeek strategy, traditional multi - factor models, and the CSI 300 Index. The results are presented in the following table:

Table 2 Backtesting Data of Various Indicators under Different Strategies and Models

Indicator	DeepSeek Strategy	Traditional Multi - factor Model	CSI 300 Index

Indicator	DeepSeek Strategy	Traditional Multi - factor Model	CSI 300 Index
Annualized Return	32.1%	18.5%	6.7%
Sharpe Ratio	1.87	0.92	0.35
Maximum Drawdown	- 12.3%	- 22.7%	- 28.4%
Industry Coverage	82%	65%	100%

In terms of the annualized return, the DeepSeek strategy reached 32.1%, which is significantly higher than 18.5% of the traditional multi - factor model and 6.7% of the CSI 300 Index, indicating that this strategy has an obvious advantage in return acquisition. The Sharpe ratio measures the additional return that an investment portfolio can obtain over the risk - free return when taking on a unit of risk. The Sharpe ratio of the DeepSeek strategy is 1.87, much higher than 0.92 of the traditional multi - factor model and 0.35 of the CSI 300 Index, suggesting that it performs outstandingly in risk - adjusted returns. Regarding the maximum drawdown, the DeepSeek strategy is - 12.3%, which is a significant reduction compared to - 22.7% of the traditional multi - factor model and - 28.4% of the CSI 300 Index, reflecting the advantage of this strategy in risk control. In terms of industry coverage, the DeepSeek strategy reaches 82%, which is higher than 65% of the traditional multi - factor model. Although it is lower than the CSI 300 Index, it still indicates that this strategy can cover a relatively large number of industries and achieve a relatively wide investment layout.

5.2 Market Structure Impact

The DeepSeek strategy has had a variety of positive impacts on the market structure.

In terms of empowering small and medium - sized investors, after a private equity fund adopted the 7B edge - side model, the strategy research and development cost decreased from 2 million yuan per year to 450,000 yuan. This enables small and medium - sized investors to obtain advanced investment strategies at a relatively low cost, reducing the investment threshold, enhancing the competitiveness of small and medium - sized investors in the market, and promoting market fairness.

In terms of capturing industry rotation, during the artificial intelligence sector market in Q4 2023, the model identified targets such as Cambricon and Sugon two weeks in advance, and the portfolio's excess return reached 18.9%. The accurate industry rotation capture ability helps investors to promptly grasp market hotspots, adjust investment portfolios, and obtain higher returns.

In terms of improving liquidity, a 1 - basis - point increase in the trading volume share of the strategy can reduce the liquidity premium of the ChiNext Index by 0.3%. This indicates that the application of this strategy helps to improve market liquidity, reduce transaction costs, and enhance market operation efficiency.

6 CASE STUDIES: INVESTMENT OPPORTUNITIES DRIVEN BY DEEPSEEK

6.1 Computing Power Infrastructure

Taking Cambricon (688256) as an example, DeepSeek - V3 analyzed the performance parameters of its Siyuan 590 chip and predicted that the market share of AI chips in 2024 would increase to 19% (7 percentage points higher than the consensus forecast of securities firms). By comparing the supply chain data of NVIDIA H20, it was found that Cambricon had an advantage in packaging and testing costs. This analysis result provides investors with a new perspective on the investment value of Cambricon. Based on this, investors can more accurately evaluate the future development potential and investment returns of Cambricon, and thus make more rational investment decisions.

6.2 Financial Technology Applications

In the case of Hundsun Technologies (600570), the 7B model conducted real - time analysis of the customer feedback of its O45 system and captured the key signal that "the customized demand of asset management institutions increased by 23%". Combined with the policy deduction ability of the R1 model, it predicted an incremental market of 1.4 billion yuan brought by the new asset management regulations. This case demonstrates the application value of the DeepSeek model in the financial technology field. Through real - time analysis of customer feedback and in - depth interpretation of policies, potential investment opportunities can be explored, providing strong support for investors' investments in the financial technology sector.

6.3 Edge - side Hardware Ecosystem

By analyzing the technology roadmap of the robotics industry, the model found that the RK3588S chip of Rockchip (603893) accounted for 18% of the BOM cost of service robots. Combined with the V3 code generation ability, it simulated the performance elasticity brought by the mass production of humanoid robots. This analysis helps investors understand the important position of Rockchip in the edge - side hardware ecosystem and its potential performance growth space, providing an important basis for investing in Rockchip.

7 Challenges and Future Directions

7.1 Existing Challenges

In terms of regulatory compliance, the black - box decision - making process of the DeepSeek model conflicts with the "Algorithmic Management Guidelines for the Securities and Futures Industry" [10]. Since the decision - making process of the model is difficult to explain intuitively, regulatory authorities face difficulties in supervising investment strategies. To solve this problem, it is necessary to develop a SHAP value visualization interpretation module to present the decision - making logic of the model in a visual way, improve the transparency of the decision - making process, and meet regulatory requirements.

Data barriers are also an important issue. The acquisition cost of unstructured data accounts for 63% of the total model cost. Unstructured data, such as news reports and social media information, although containing rich market information, is difficult to acquire and process. Currently, it is necessary to explore federated learning solutions to carry out data collaboration and model training without disclosing the original data, reducing the data acquisition cost while protecting data privacy.

7.2 Evolution Path

In the future, the DeepSeek model has broad prospects for development in the field of intelligent stock selection in the A-share market. Its evolution path will revolve around three key directions: multimodal integration, real-time leapfrogging, and compliance innovation. By doing so, it will continuously enhance the model's performance and market adaptability, and further promote the intelligent transformation of A-share investment.

In terms of multimodal integration, there is great potential in accessing the DeepSeek-V3 visual module to analyze the industrial chain map. Take the photovoltaic industry as an example. By obtaining the workshop monitoring data of photovoltaic enterprises, the model can intuitively understand the operating status of enterprise production equipment, the efficiency of the production process, and the real-time situation of product quality.

Real-time leapfrogging is an important direction for the development of the DeepSeek model. The adoption of MTP (Multi-Token Prediction) technology can significantly improve the generation speed of trading signals, making it reach the microsecond level. In the current rapidly changing financial market, trading opportunities are fleeting. For example, in high-frequency trading scenarios, when there is sudden news or abnormal market fluctuations, the microsecond-level trading signal generation speed allows investors to react quickly.

Compliance innovation is a crucial aspect that must be emphasized in the development process of the DeepSeek model. Developing a "regulatory sandbox" version and embedding intelligent contracts for investor suitability management are important measures to achieve compliant development. The "regulatory sandbox" provides a safe testing environment for new investment strategies and models. In this environment, the model can operate under simulated market conditions while being monitored and evaluated in real time by regulatory authorities.

With the continuous deepening of multimodal integration, the gradual realization of real-time leapfrogging, and the continuous advancement of compliance innovation, the DeepSeek model will play an even more powerful role in the field of intelligent stock selection in the A-share market. It can not only provide investors with more accurate and efficient investment decision-making support, but also promote the intelligent process of the entire capital market, facilitate the optimal allocation of financial resources, and help the Chinese capital market move towards a stage of higher-quality development.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Eugene F, Fama, Kenneth R, French. Fama - French Three - Factor Model and Its Application. 1993. Chinese Software Developer Network. Retrieved from https://blog.csdn.net/hzk427/article/details/104187847

[2] Liu Yaning. ChatGPT Consumes Over 500,000 kWh of Electricity per Day. How Energy - Intensive Are AI Large - Scale Models? Titanium Media. 2024. Retrieved from https://baijiahao.baidu.com/s?id = 1793597173423030217&wfr = spider&for = pc

[3] Qiao Nan. Decoding the Innovation Path of DeepSeek: The Evolution Roadmap of Three Generations of Models. 2025. Retrieved from https://news.sina.cn/ai/2025 - 02 - 09/detail - ineiwqye4694847.d.html

[4] Liao Hailong. DeepSeek Large - Scale Model: Technical Analysis and Development Prospect. Journal of Computer Science and Electrical Engineering, 2025, 7(1): 33-37. DOI: https://doi.org/10.61784/jcsee3035

[5] Lieyan Benniu. DeepSeek: Algorithm Innovation Leads to an AI Computing Power Boom and Usher in a New Era of Intelligent Technology. 2025. Retrieved from http://mp.weixin.qq.com/s?__biz = Mzg2MTE3NDA2Mg==&mid = 2247494898&idx = 1&sn = b99504080dcd03e5a6809136486d5bde&scene = 0

[6] Wang Houdong. The Important Impact of DeepSeek on the Global Development of Artificial Intelligence. 2025. Retrieved from http://mp.weixin.qq.com/s?__biz = MjM5MjgzMzg1Nw==&mid = 2453991839&idx = 1&sn = 74a9185fb47c86c8deb4939df9b81cd5&scene = 0

[7] Demon Akana. Introduction to the Technical Principles of Deepseek's MLA Technology. Chinese Software Developer Network. 2025. Retrieved from https://blog.csdn.net/bestpasu/article/details/145539423

[8] Baitai Laoren. DeepSeekMoE Architecture. Chinese Software Developer Network. 2024. Retrieved from https://blog.csdn.net/weixin_41429382/article/details/144701990

[9] Zhihong Shao, Peiyi Wang, Qihao Zhu, et al. DeepSeekMath: Pushing the Limits of Mathematical Reasoning in Open Language Models. 2024. DOI: https://doi.org/10.48550/arXiv.2402.03300. Retrieved from https://arxiv.org/pdf/2402.03300

[10] Le Tutu. In - Depth Analysis of Other Potential Defects of DeepSeek. Eastmoney Wealth Channel. 2025. Retrieved from https://caifuhao.eastmoney.com/news/20250202221538225857610