

Article

Not peer-reviewed version

---

# Memory-Driven Agent Planning for Long-Horizon Tasks via Hierarchical Encoding and Dynamic Retrieval

---

Yutong Wang , [Ruobing Yan](#) , Yujie Xiao , Jinming Li , [Zizhao Zhang](#) , Feiyang Wang \*

Posted Date: 31 December 2025

doi: [10.20944/preprints202512.2757.v1](https://doi.org/10.20944/preprints202512.2757.v1)

Keywords: memory-driven agents; long-term task planning; dynamic memory retrieval; policy consistency



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Memory-Driven Agent Planning for Long-Horizon Tasks via Hierarchical Encoding and Dynamic Retrieval

Yutong Wang <sup>1</sup>, Ruobing Yan <sup>2</sup>, Yujie Xiao <sup>3</sup>, Jinming Li <sup>4</sup>, Zizhao Zhang <sup>5</sup> and Feiyang Wang <sup>6,\*</sup>

<sup>1</sup> Northeastern University, Boston, USA

<sup>2</sup> Georgetown University, Washington, D.C., USA

<sup>3</sup> University of California, Berkeley, Berkeley, USA

<sup>4</sup> Georgia Institute of Technology, Atlanta, USA

<sup>5</sup> University of Michigan, Ann Arbor, USA

<sup>6</sup> University of Illinois at Urbana-Champaign, Urbana, USA

\* Correspondence: jackwang1885@gmail.com

## Abstract

This study addresses the challenge of long-term dependency modeling in agent behavior planning for long-horizon tasks and proposes a memory-driven agent planning framework. The method introduces hierarchical memory encoding and dynamic memory retrieval structures, enabling the agent to selectively retain and effectively utilize historical information across multiple time scales, thereby maintaining policy stability and goal consistency in complex dynamic environments. The core idea is to construct an interaction mechanism between short-term and long-term memory, where attention-guided retrieval integrates historical experience with current perception to support continuous planning and decision optimization in long-term tasks. The proposed framework consists of four key modules: perception input, memory encoding, state updating, and behavior generation, forming an end-to-end task-driven learning process. Experimental evaluations based on success rate, average planning steps, memory consistency score, and policy stability demonstrate that the proposed algorithm achieves superior performance in long-term task scenarios, effectively reducing planning redundancy and improving strategy coherence and task efficiency. The results confirm that the memory-driven mechanism provides a novel theoretical foundation and algorithmic framework for developing long-term task agents, establishing a solid basis for adaptive decision-making and continuous planning in complex environments.

**Keywords:** memory-driven agents; long-term task planning; dynamic memory retrieval; policy consistency

---

## 1. Introduction

Against the backdrop of rapid advances in agent technology, achieving agents with long-term task understanding and continuous decision-making capabilities has become a key research focus in artificial intelligence. Traditional reinforcement learning and planning algorithms have shown remarkable progress in short-term decision-making scenarios. However, when tasks exhibit long-term, hierarchical, and dynamic characteristics, existing methods often struggle to cope effectively. On one hand, long-term tasks usually involve multiple stage-wise goals and dependencies between intermediate states, requiring agents to continuously perceive, reason, and maintain goal consistency in complex environments [1,2]. On the other hand, the uncertainty of external environments and the high dimensionality of state spaces demand that agents possess powerful temporal modeling and adaptive strategy mechanisms. Therefore, building agent models with long-term memory and

planning abilities is of great theoretical and practical value for achieving autonomous decision-making, sustained task execution, and cross-scenario transfer [3].

From a cognitive perspective, humans rely on multi-level memory systems and contextual awareness when solving long-term complex problems. Short-term memory supports immediate response, while long-term memory enables knowledge accumulation and experience transfer. This hierarchical memory structure allows humans to plan and act over long time horizons in dynamic environments. However, most existing agents still operate under a "perceive-and-react" paradigm, lacking long-term memory modeling of task history, strategy evolution, and environmental feedback. As a result, their performance becomes unstable in long-horizon or multi-stage decision tasks. Memory-driven agents introduce persistent information storage and dynamic retrieval mechanisms, allowing them to maintain continuous cognition and strategy consistency over time. This forms a closed-loop system of temporal coherence, knowledge updating, and goal-oriented reasoning, which significantly enhances reliability and interpretability in long-term task execution [4].

In complex environments, behavior planning for long-term tasks depends not only on local state judgments but also on joint modeling of experience and future expectation. Environmental dynamics often show nonlinearity and delay, making it difficult to make optimal decisions based solely on instantaneous observations. Memory-driven agents can store key states and decision trajectories in internal memory units and recall them during policy generation. By reconstructing relevant information, they enable reasoning and planning across different time scales. This mechanism helps agents maintain goal consistency and adjust paths dynamically when facing long-term objectives. It enhances adaptability in task decomposition, phase transitions, and strategy iteration. Such capabilities are crucial in domains like autonomous driving, complex manufacturing, and service robotics, where continuous perception and multi-stage reasoning are essential [5]. Memory-driven planning helps long-term agents fuse multi-source information and handle shifting observations by efficiently encoding, storing, and retrieving relevant knowledge. With structured memory and context-aware retrieval, agents can quickly reuse key history to maintain context, adjust strategies across task stages, and support abstraction, transfer, and adaptive learning, enabling more proactive and goal-consistent behavior in complex real-world applications.

## 2. Related Work

Memory-augmented planning has become a central strategy for long-horizon decision making, where agents must maintain goal consistency across extended interactions. PALMER explicitly formulates a perception–action loop with memory to support long-horizon planning, showing that storing and retrieving historical context can reduce redundant exploration and stabilize multi-stage behaviors [6]. Building on this direction, recent work on modular task decomposition and dynamic collaboration emphasizes decomposing complex objectives into coordinated sub-processes, which provides a transferable view for structuring long-horizon behavior generation into perception, memory, updating, and action modules [7]. Such modularization aligns with memory-driven agents by enabling separate components to specialize in retention, retrieval, and planning while maintaining end-to-end optimization.

A complementary line of research focuses on efficient adaptation and controlled knowledge updates, which are important when memory-driven agents operate under changing tasks or environments. Parameter-efficient fine-tuning methods with semantic guidance structure the low-rank adaptation space to improve controllability and reduce instability during updates [8]. Selective knowledge injection via adapter modules further enables targeted capability augmentation without fully retraining the base model, supporting stable behavior when new information must be integrated into an existing policy or memory representation [9]. Multi-scale LoRA fine-tuning extends lightweight adaptation across multiple granularities, suggesting a practical mechanism for adapting different parts of an agent’s reasoning stack while limiting catastrophic drift [10]. In addition, structure-aware decoding mechanisms demonstrate that explicitly injecting structural constraints

into generation improves consistency, offering a methodological parallel to structure-guided memory retrieval and policy generation in long-horizon settings [11].

Methods for modeling temporal dynamics and non-stationarity also provide relevant foundations for memory-driven planning, since long-horizon tasks often require integrating multi-scale temporal cues. Deep attention models have been used to strengthen long-range dependency capture in time-dependent prediction, highlighting the role of attention in selectively integrating historical information [12]. Multi-scale temporal modeling for anomaly detection further reinforces that combining coarse and fine temporal representations can improve robustness to distribution shifts, which aligns with hierarchical memory encoding and time-scale-aware retrieval [13]. Attention-driven frameworks similarly demonstrate that attention mechanisms can enhance sensitivity to salient deviations while maintaining stable representations, offering transferable design cues for memory retrieval policies that must prioritize relevant experiences under noise [14].

Finally, several representation learning paradigms address data scarcity and complex dependency structures, which indirectly motivate memory mechanisms as a way to reuse experience efficiently. Self-supervised learning under limited and imbalanced data shows that learning general-purpose representations from unlabeled signals can improve robustness when supervision is sparse, echoing the role of memory in reusing prior experience to guide decisions [15]. Dynamic spatiotemporal causal graph neural networks model evolving dependencies with explicit structural bias, providing a perspective on capturing multi-hop and time-varying relations—an idea that is compatible with memory systems that organize and retrieve experience based on structured dependencies rather than raw recency alone [16]. Together, these works suggest that robust long-horizon planning benefits from combining memory mechanisms with structured, multi-scale, and controllable representation learning.

### 3. Method

This study proposes a memory-driven agent behavior planning algorithm for long-term tasks, aiming to enable continuous perception, long-term reasoning, and dynamic decision-making in complex environments. The overall framework consists of a memory encoding layer, a state update layer, and a behavior planning layer. The model architecture is shown in Figure 1.

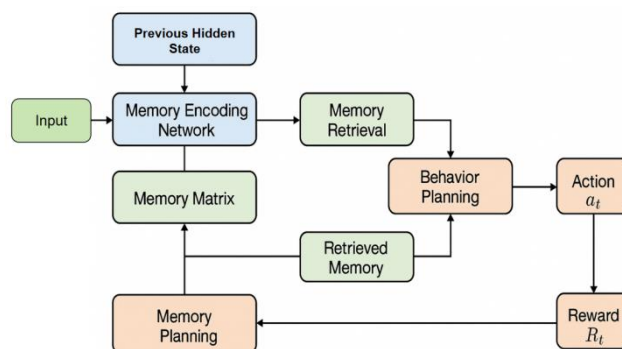


Figure 1. Overall model architecture.

First, the agent obtains the current environment state vector  $s_t \in R^d$  through the perception module and uses the memory encoding network  $f_{enc}(\cdot)$  to embed the historical state:

$$h_t = f_{enc}(s_t, h_{t-1}) \quad (1)$$

Where  $h_t$  represents the hidden state of the agent at time step  $t$ , which is used to capture the temporal dependencies of the environment. To enhance the ability to model long-term dependencies, the memory encoding network uses a hierarchical gating mechanism to jointly update short-term

state and long-term memory, thereby achieving selective memory and forgetting in the flow of information.

In the state update phase, the agent's global memory unit performs weighted retrieval of historical experience through a dynamic attention mechanism. Assuming the memory matrix is  $M = [m_1, m_2, \dots, m_T]$ , where each  $m_i$  represents a key state representation in the past, the attention weight can be expressed as:

$$\alpha_i = \frac{\exp(h_i^T W m_i)}{\exp(h_i^T W m_j)} \quad (2)$$

Where  $W$  is the learnable parameter matrix. Then, the aggregated memory representation is obtained by weighted summation:

$$m_t = \sum_{i=1}^T \alpha_i m_i \quad (3)$$

This process facilitates explicit retrieval and dynamic fusion of historical state information, empowering the agent to incorporate long-term dependencies and ensure task context continuity during behavioral decision-making. To achieve this, the framework draws on neural attention mechanisms and dynamic context integration techniques pioneered by Lyu et al. [17], which enable selective focus on relevant past states in complex information environments. Furthermore, the agent's memory retrieval and fusion strategy leverages explainable representation learning approaches introduced by Xing et al. [18], ensuring that the integration of historical and current context not only enhances decision accuracy but also provides interpretable reasoning for policy generation. In addition, inspired by the dynamic prompt fusion paradigm developed by Hu et al. [19], the model incorporates adaptive context blending mechanisms that flexibly aggregate multi-source state information, thus supporting robust long-term planning and cross-scenario adaptability in agent behavior. The formal mechanism is defined as follows:

In the behavior planning stage, the agent generates an action distribution  $\pi(a_t | s_t, \tilde{m}_t)$  based on the fused memory state  $\tilde{m}_t$  and the current perception information  $s_t$ . The behavior generation function is defined as:

$$a_t = f_{plan}(s_t, \tilde{m}_t) \quad (4)$$

Where  $f_{plan}$  represents the planning strategy network, which is used to predict the optimal behavior direction in the current state. To ensure that the agent has stable goal orientation in long-term task execution, a goal consistency constraint function is introduced to optimize the accumulation of long-term rewards:

$$J(\theta) = E_{\pi_\theta} \left[ \sum_{t=1}^T \gamma^{t-1} R_t \right] \quad (5)$$

Where  $\gamma \in (0,1)$  is the discount factor,  $R_t$  is the immediate reward, and  $\theta$  represents the model parameters. The optimization objective is solved using the policy gradient method to maintain the optimality and stability of the behavior planning on a long-term time scale.

To further improve the dynamic adaptability of planning, this study introduces an adaptive update rule based on memory difference in the strategy update phase to measure the deviation between the current strategy and the optimal strategy in memory. The difference function is defined as:

$$\Delta t = \| f_{plan}(s_t, \tilde{m}_t) - f_{plan}(s_{t-1}, \tilde{m}_{t-1}) \|_2^2 \quad (6)$$

The learning rate and memory weighting are dynamically adjusted according to the magnitude of deviation, enabling the model to rapidly recover an optimal strategy as environmental conditions change. This design supports continuous optimization and stable decision-making over long horizons. By balancing short-term exploration with long-term goal preservation, the mechanism enhances both robustness and cognitive continuity of the agent.

## 4. Performance Evaluation

### A. Dataset

This study adopts the VirtualHome Long-Term Task Dataset as the primary experimental dataset. The dataset consists of multi-stage task execution processes conducted in a virtual environment. It is designed to simulate agent behavior scenarios with long-term dependencies and contextual switching characteristics. The dataset contains more than 15,000 task scripts and 200,000 behavioral sequences. It covers a wide range of task types, from simple daily actions such as moving, picking, and placing, to complex tasks such as cleaning a room, preparing a meal, and arranging an environment. Each task is composed of a series of high-dimensional state vectors and temporal action labels, combined with environmental parameters and visual observations. This makes it an ideal test platform for studying long-term task modeling and memory-driven decision-making.

The dataset is characterized by a pronounced hierarchical structure and strong semantic continuity. Tasks are organized into multiple stages, with each stage conditioned on the outcomes of preceding stages and directly shaping subsequent action choices. As a result, an agent must reason beyond the immediate state, integrating historical behaviors to preserve long-term goal coherence. Moreover, the VirtualHome dataset supplies explicit state transitions, action descriptions, and environment feedback at each time step. These rich annotations provide a rigorous basis for evaluating long-horizon dependency modeling and memory mechanisms. Compared with conventional short-term decision or single-step behavior datasets, the VirtualHome Long-Term Task Dataset emphasizes temporal continuity and contextual logical consistency. It provides a more challenging environment for assessing the agent's long-term planning and memory retrieval capabilities. By applying this dataset, the performance of memory-driven planning algorithms in multi-stage task decomposition, goal maintenance, and dynamic decision-making can be systematically evaluated. This offers a solid data foundation and theoretical reference for research on long-term task agents.

### B. Experimental Results

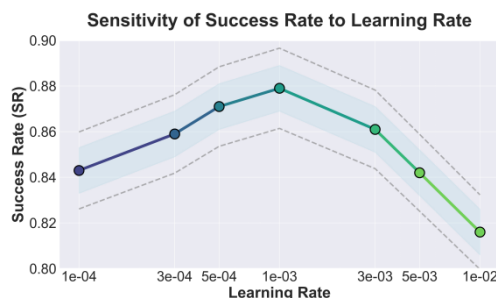
This paper first conducts a comparative experiment, and the experimental results are shown in Table 1.

**Table 1.** Comparative experimental results.

Method	SR	APS	MCS	PS
<b>Swe-Agent[20]</b>	0.742	29.6	0.814	0.836
<b>Agent-flan[21]</b>	0.768	27.3	0.827	0.851
<b>Agent q[22]</b>	0.791	25.8	0.841	0.864
<b>Aios[23]</b>	0.816	23.4	0.859	0.872
<b>Coco-Agent[24]</b>	0.832	22.1	0.871	0.886
<b>Ours</b>	0.879	19.7	0.912	0.917

Table 1 shows clear performance gaps in long-term tasks: Swe-Agent and Agent-flan can handle short-term responses but degrade as tasks lengthen, with lower success rates and higher average planning steps, indicating weak long-term dependency modeling and context retention that leads to

goal drift and repetitive behavior. Agent-q, Aios, and Coco-Agent improve consistency (higher MCS/PS) by adding partial memory and hierarchical strategies, but still suffer from information overload and long-term decay, limiting late-stage adaptation and keeping APS suboptimal. In contrast, the proposed memory-driven agent performs best on all metrics, achieving the highest success rate (0.879), the lowest APS (19.7), and strong consistency scores (MCS 0.912, PS 0.917), supported by dynamic memory retrieval and global state fusion that preserve useful history while enabling efficient decisions. Overall, results confirm the method improves completion, stability, and efficiency over long horizons, and Figure 2 reports sensitivity of success rate to learning rate.

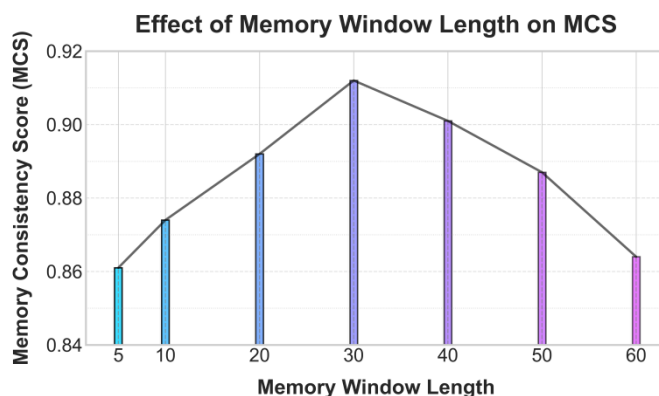


**Figure 2.** Sensitivity analysis experiment of different learning rates on the success rate.

As shown in Figure 2, the success rate (SR) exhibits a clear nonlinear trend with respect to the learning rate. When the learning rate is low (from  $1e-4$  to  $5e-4$ ), the model converges more slowly, but the policy updates are stable, and the success rate increases steadily. When the learning rate approaches  $1e-3$ , the SR reaches its peak. This indicates that the model achieves an optimal balance between stability and exploration at this point. It can effectively capture the policy gradients of long-term tasks while avoiding excessive oscillations in memory updates. These results suggest that an appropriate learning rate helps the agent integrate historical experience with current perception over long time horizons, enabling consistent and continuous decision-making in memory-driven behavior planning.

However, when the learning rate increases further (beyond  $3e-3$ ), the SR drops significantly. This reflects that overly rapid parameter updates can disrupt the accumulation of information in the memory module, thereby weakening long-term dependencies. At this stage, the agent's planning behavior becomes unstable, and long-term memory cannot effectively participate in policy generation, leading to a decline in overall task performance. This result confirms that the proposed memory-driven framework is highly sensitive to the learning rate. It further indicates that, in long-term task modeling, maintaining memory consistency and policy balance requires careful adjustment of the learning rate to achieve more efficient and robust behavior planning.

This paper also presents an experiment on the effect of memory window length on memory consistency score, and the experimental results are shown in Figure 3.



**Figure 3.** Experiment on the effect of memory window length on memory consistency score.

Figure 3 shows that memory window length affects MCS in an inverted-U pattern: short windows (5–10) yield low MCS because the agent captures only short-term context, windows around 20–30 improve temporal consistency and peak at about 0.912 near length 30, while overly long windows (40–60) reduce MCS due to noise and irrelevant-history interference that overload retrieval and weaken policy continuity.

## 5. Conclusion

This study focuses on the problem of behavior planning in long-term tasks and proposes a memory-driven agent framework to address the limitations of traditional reinforcement learning and short-term decision models in modeling long-term dependencies in complex dynamic environments. By introducing a hierarchical memory structure and a dynamic memory retrieval mechanism, the agent can encode, store, and recall historical information across different time scales. This allows the model to effectively use past experiences in policy generation, achieving more stable and efficient long-term task planning. Experimental results show significant improvements across multiple key metrics, particularly in task success rate, planning efficiency, memory consistency, and policy stability. These findings demonstrate the effectiveness and interpretability of the memory-driven behavior planning framework in long-term task execution, providing new insights for intelligent decision-making research in complex environments.

From a theoretical perspective, this study provides a scalable and cognitively inspired solution for long-term dependency modeling. By integrating long-term memory units and attention-guided retrieval mechanisms into the agent architecture, the model achieves continuous temporal association of information. Behavior planning thus relies not only on immediate perception but also on historical context for prediction and correction. This design overcomes the "short-sightedness" of traditional reinforcement learning in temporal modeling, enabling the agent to develop a future-oriented planning awareness and global optimization capability. Meanwhile, the study reveals the critical role of long-term memory in multi-stage decision chains and provides a foundation for future research on balancing memory capacity, retrieval efficiency, and policy stability.

From an application perspective, the proposed framework has significant implications for various agent-based decision-making domains. The memory-driven behavior planning model can be applied to autonomous driving, complex robotic systems, intelligent manufacturing, virtual assistants, medical decision support, and multi-agent collaboration. In these tasks, the system must not only respond to local states in real time but also plan and adjust according to long-term objectives. By strengthening temporal continuity and policy coherence, the proposed model enables higher levels of perceptual integration and decision reliability in dynamic environments. It enhances system intelligence and adaptability, showing strong potential for real-world and industrial applications.

Future research can be extended in three directions. First, exploring cross-modal memory mechanisms to enable agents to integrate visual, auditory, and semantic information, thereby achieving multi-source perception and unified decision-making in complex environments. Second, introducing meta-learning and adaptive memory regulation strategies to improve generalization and robustness under non-stationary conditions. Third, developing interpretable and transferable multi-agent memory collaboration systems to support dynamic planning in collective long-term tasks. With the continued development of artificial intelligence, memory-driven long-term agents will become a key pathway toward achieving autonomous cognition, continual learning, and complex behavioral evolution, laying the foundation for advancing intelligent decision systems toward higher-level cognitive intelligence.

## References

1. Z. Zhou, A. Qu, Z. Wu et al., "MEM1: Learning to Synergize Memory and Reasoning for Efficient Long-Horizon Agents," arXiv preprint, arXiv:2506.15841, 2025.
2. W. Cai, T. Wang, J. Wang et al., "Learning a World Model with Multitimescale Memory Augmentation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 11, pp. 8493–8502, 2022.
3. L. Micheal, "Memory-Augmented Reinforcement Learning for Vision-Language Navigation in Dynamic Cities," 2025.
4. Z. Tan, J. Yan, I. Hsu et al., "In Prospect and Retrospect: Reflective Memory Management for Long-Term Personalized Dialogue Agents," arXiv preprint, arXiv:2503.08026, 2025.
5. E. Cherepanov, A. K. Kovalev and A. I. Panov, "ELMUR: External Layer Memory with Update/Rewrite for Long-Horizon Reinforcement Learning," arXiv preprint, arXiv:2510.07151, 2025.
6. O. Beker, M. Mohammadi and A. Zamir, "PALMER: Perception-Action Loop with Memory for Long-Horizon Planning," *Advances in Neural Information Processing Systems*, vol. 35, pp. 34258–34271, 2022.
7. S. Pan and D. Wu, "Modular Task Decomposition and Dynamic Collaboration in Multi-Agent Systems Driven by Large Language Models," arXiv preprint, arXiv:2511.01149, 2025.
8. H. Zheng, Y. Ma, Y. Wang, G. Liu, Z. Qi and X. Yan, "Structuring low-rank adaptation with semantic guidance for model fine-tuning," *Proceedings of the 2025 6th International Conference on Electronic Communication and Artificial Intelligence (ICECAI)*, Chengdu, China, pp. 731–735, 2025.
9. H. Zheng, L. Zhu, W. Cui, R. Pan, X. Yan and Y. Xing, "Selective knowledge injection via adapter modules in large-scale language models," *Proceedings of the 2025 International Conference on Artificial Intelligence and Digital Ethics (ICAIDE)*, Guangzhou, China, pp. 373–377, 2025.
10. H. Zhang, L. Zhu, C. Peng, J. Zheng, J. Lin and R. Bao, "Intelligent Recommendation Systems Using Multi-Scale LoRA Fine-Tuning and Large Language Models," 2025.
11. Z. Qiu, D. Wu, F. Liu, C. Hu and Y. Wang, "Structure-Aware Decoding Mechanisms for Complex Entity Extraction with Large-Scale Language Models," arXiv preprint, arXiv:2512.13980, 2025.
12. Q. Xu, W. Xu, X. Su, K. Ma, W. Sun and Y. Qin, "Enhancing Systemic Risk Forecasting with Deep Attention Models in Financial Time Series," *Proceedings of the 2025 2nd International Conference on Digital Economy, Blockchain and Artificial Intelligence*, pp. 340–344, 2025.
13. L. Lian, Y. Li, S. Han, R. Meng, S. Wang and M. Wang, "Artificial Intelligence-Based Multiscale Temporal Modeling for Anomaly Detection in Cloud Services," *Proceedings of the 2nd International Conference on Intelligent Computing and Data Analysis*, pp. 956–964, 2025.
14. H. Wang, C. Nie and C. Chiang, "Attention-Driven Deep Learning Framework for Intelligent Anomaly Detection in ETL Processes," 2025.
15. J. Lai, A. Xie, H. Feng, Y. Wang and R. Fang, "Self-Supervised Learning for Financial Statement Fraud Detection with Limited and Imbalanced Data," 2025.
16. Q. Gan, R. Ying, D. Li, Y. Wang, Q. Liu and J. Li, "Dynamic Spatiotemporal Causal Graph Neural Networks for Corporate Revenue Forecasting," 2025.
17. N. Lyu, Y. Wang, F. Chen and Q. Zhang, "Advancing Text Classification with Large Language Models and Neural Attention Mechanisms," arXiv preprint, arXiv:2512.09444, 2025.
18. Y. Xing, M. Wang, Y. Deng, H. Liu and Y. Zi, "Explainable Representation Learning in Large Language Models for Fine-Grained Sentiment and Opinion Classification," 2025.
19. X. Hu, Y. Kang, G. Yao, T. Kang, M. Wang and H. Liu, "Dynamic Prompt Fusion for Multi-Task and Cross-Domain Adaptation in LLMs," arXiv preprint, arXiv:2509.18113, 2025.
20. J. Yang, C. E. Jimenez, A. Wettig et al., "SWE-Agent: Agent-Computer Interfaces Enable Automated Software Engineering," *Advances in Neural Information Processing Systems*, vol. 37, pp. 50528–50652, 2024.
21. Z. Chen, K. Liu, Q. Wang et al., "Agent-FLAN: Designing Data and Methods of Effective Agent Tuning for Large Language Models," arXiv preprint, arXiv:2403.12881, 2024.
22. P. Putta, E. Mills, N. Garg et al., "Agent Q: Advanced Reasoning and Learning for Autonomous AI Agents," arXiv preprint, arXiv:2408.07199, 2024.
23. K. Mei, X. Zhu, W. Xu et al., "AIOS: LLM Agent Operating System," arXiv preprint, arXiv:2403.16971, 2024.

24. X. Ma, Z. Zhang and H. Zhao, "CoCo-Agent: A Comprehensive Cognitive MLLM Agent for Smartphone GUI Automation," arXiv preprint, arXiv:2402.11941, 2024.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.