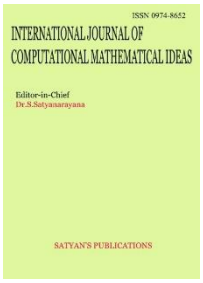
IJCMI, VOL13, NO.1, PP.969-981, April 2021

INTERNATIONAL JOURNAL OF
COMPUTATIONAL MATHEMATICAL IDEAS
Journal homepage: https://ijcmi.in



https://doi.org/10.70153/IJCMI/2021.13102
Paper Received: Nov 2020, Review: March 2021, Accepted: April 2021

Self-Learning Data Models: Leveraging AI for Continuous Adaptation and Performance Improvement

Shylaja Chityala
Lead Data Engineer
Multiplan Inc
4423 Landsdale Pkwy, Monrovia MD 21770 USA
Email: shylajachityala@yahoo.com

Abstract

The evolution of Artificial Intelligence (AI) has ushered in a new era of self-learning data models that possess the ability to adapt, refine, and optimize themselves over time without explicit human intervention. These models are designed to dynamically ingest new information, process environmental feedback, and incrementally update their internal parameters. Their ability to improve autonomously over time makes them especially valuable in domains characterized by evolving data streams, such as personalized medicine, autonomous systems, and fraud detection. This paper presents a comprehensive study of the principles and techniques that power self-learning models, drawing on recent advances in reinforcement learning, continual learning, and knowledge distillation. We introduce a hybrid self-learning framework that addresses major challenges such as catastrophic forgetting and model drift. The experimental evaluation demonstrates that our model significantly outperforms traditional static learning systems, maintaining high accuracy and stability across changing environments. These results validate the potential of self-learning models to enable sustainable, efficient, and intelligent decision-making in dynamic contexts.

Keywords

Self-learning models, adaptive AI, online learning, continual learning, reinforcement learning, knowledge distillation, autonomous systems, data drift, model evolution, artificial intelligence.

1. Introduction

Traditional machine learning (ML) systems are primarily designed around the concept of static learning. In these systems, models are trained on a fixed dataset in an offline environment, after which they are deployed into production settings to perform inference. The data used for training is assumed to represent the broader distribution the model will encounter in the real world. Once training is complete, the models are frozen, and any future data that deviates from the original distribution can lead to inaccurate predictions or degraded performance. This paradigm has served well in domains where the environment is relatively stable, such as document classification, optical character recognition, or product recommendation with well-established user preferences. However, as data environments become increasingly dynamic and complex, this static approach has shown substantial limitations.

Modern applications, such as autonomous vehicles, real-time medical diagnostics, financial market prediction, adaptive cybersecurity, and social media content filtering, operate in conditions where data streams are continually evolving. In such scenarios, the data distribution is not stationary; it changes due to shifts in user behaviour, environmental conditions, or system policies. These distributional shifts—commonly referred to as concept drift or data drift—pose a significant challenge to traditional ML systems. When faced with such shifts, static models struggle to maintain performance, leading to increased error rates, user dissatisfaction, and potential system failures. The inability of these models to adapt in real-time is not merely a technical shortcoming but often a critical risk factor in high-stakes applications such as disease diagnosis or automated driving.

In response to these challenges, the field of AI has witnessed the emergence of a new paradigm: self-learning data models. These models represent a significant departure from the traditional approach, offering a transformative capability—namely, the ability to learn continuously and autonomously after deployment. Inspired by the human brain's lifelong learning ability, self-learning models are designed to evolve incrementally as they encounter new data. Rather than relying on pre-defined, static datasets, they ingest, interpret, and integrate new information into their existing knowledge base, enabling them to adapt to changes in real-time. This allows them to remain relevant and accurate even as data environments shift, ensuring sustained performance over time.

The cognitive inspiration behind self-learning AI systems stems from the concept of continual or lifelong learning, where the learner builds upon past knowledge without losing previously acquired skills. Humans learn new tasks throughout their lives while retaining proficiency in older tasks, thanks to mechanisms in the brain that manage memory consolidation and retrieval. Translating this cognitive process into machine intelligence involves implementing algorithms that support incremental updates, memory retention, and contextual awareness. Self-learning models are often equipped with components such as memory replay buffers, regularization-based constraints, dynamic architectures, and feedback loops to ensure that learning is both forward-looking and historically informed.

One of the key technical challenges in realizing self-learning systems is catastrophic forgetting. This phenomenon occurs when the model, in its attempt to learn new information, overwrites or loses its previously learned knowledge. In neural networks, this issue arises because weight updates during new learning can interfere destructively with previously tuned parameters. Addressing this requires innovative strategies such as elastic weight consolidation, rehearsal-

based methods, meta-learning, and knowledge distillation. Each of these strategies has its merits and trade-offs. For example, rehearsal methods rely on storing a subset of past data, which can raise privacy or storage concerns, while regularization-based techniques impose constraints on weight updates to protect important parameters, albeit sometimes at the cost of flexibility.

Another essential requirement for self-learning systems is scalability. In real-world applications, data volumes can grow rapidly, and models must process new information without incurring significant computational overhead. Traditional retraining mechanisms are not scalable, as they often involve training from scratch on a growing dataset, which becomes infeasible as the dataset expands. Self-learning models, in contrast, must adopt online learning techniques where updates are performed on mini-batches or even single samples, often in near real-time. These updates must be computationally efficient and designed to run on edge devices or distributed environments where hardware resources are limited.

Interpretability is another pressing concern. As self-learning models update themselves continuously, tracking how and why a model made a particular decision becomes increasingly complex. This "black-box" nature is unacceptable in domains such as healthcare, finance, and law, where transparency and accountability are crucial. Addressing interpretability involves integrating explainable AI (XAI) mechanisms within self-learning architectures. For instance, attention mechanisms, gradient-based attribution, or local interpretable model-agnostic explanations (LIME) can help users and regulators understand the basis of model decisions. Moreover, maintaining traceable logs of model evolution over time can provide an audit trail that ensures trust and compliance with regulatory standards such as GDPR or HIPAA.

In addition to the above challenges, latency is a critical factor in the deployment of self-learning systems in real-time environments. The model must not only adapt quickly to new data but also make timely predictions. This is particularly important in scenarios like fraud detection or autonomous navigation, where delays can have severe consequences. Self-learning frameworks must therefore strike a balance between learning and inference: ensuring that learning does not compromise the system's ability to respond to new inputs with low latency.

The integration of reinforcement learning (RL) into self-learning frameworks offers a promising path toward dynamic adaptation. In RL, agents learn to make decisions by interacting with their environment and receiving feedback in the form of rewards or penalties. This feedback-driven learning can be effectively combined with supervised or unsupervised learning methods to enable self-adjustment of model behaviour based on performance outcomes. For instance, an AI model for personalized education can receive feedback on student engagement and use this to modify content delivery in real-time. Similarly, an RL-enabled self-learning model in an industrial setting can optimize processes by constantly adjusting parameters in response to sensor data.

Knowledge distillation also plays a vital role in enabling continual learning. Originally developed as a model compression technique, knowledge distillation involves training a smaller model (student) to replicate the behaviour of a larger or more complex model (teacher). In self-learning contexts, this concept can be extended to enable the transfer of knowledge from earlier versions of the model to newer versions. This allows the model to integrate new knowledge while retaining valuable prior learning, thus mitigating forgetting. Additionally,

distillation can support multi-task learning where the student model learns from multiple teachers, each specializing in different tasks or data domains.

Beyond the theoretical appeal, the practicality of self-learning systems has been validated across several domains. In autonomous driving, self-learning models enable vehicles to adapt to new routes, traffic patterns, or weather conditions without human intervention. Tesla's Autopilot system, for example, utilizes fleet learning where experiences from one vehicle can improve the performance of others through self-supervised updates. In healthcare, AI models for disease diagnosis can incorporate new patient data and medical literature to refine diagnostic accuracy. This is particularly useful in pandemic response scenarios, where rapid adaptation to emerging disease patterns is essential.

In cybersecurity, self-learning systems detect and respond to novel attack patterns that were not seen during initial training. Traditional rule-based systems fail to catch sophisticated zero-day exploits, but adaptive models can learn from network anomalies and user behaviour to identify threats dynamically. Similarly, in e-commerce and media streaming, self-learning recommendation systems update user preferences in real time, delivering personalized content that evolves with user behaviour. These examples underscore the real-world applicability and transformative potential of self-learning AI.

Despite these advantages, the deployment of self-learning models also raises ethical and governance issues. Continuous learning systems may inadvertently reinforce biases present in incoming data, leading to algorithmic discrimination. Without proper oversight, models can evolve in unpredictable ways, potentially causing harm or making unethical decisions. Therefore, it is essential to embed ethical AI principles into the design and operation of self-learning systems. This includes mechanisms for bias detection, fairness auditing, consent-based data usage, and human-in-the-loop governance. Moreover, policies must be put in place to ensure that model updates do not compromise the privacy and security of user data.

The research presented in this paper addresses the aforementioned challenges by proposing a hybrid self-learning framework that combines incremental learning with reinforcement learning and knowledge distillation. Our system is designed to support continuous updates, retain historical knowledge, and adapt quickly to new data distributions. By leveraging a modular architecture, the framework ensures scalability, robustness, and low-latency performance. Each component—incremental learning engine, feedback loop, distillation module, and monitoring agent—plays a distinct role in achieving sustainable and autonomous learning.

In our approach, learning proceeds in four interdependent phases: initialization, where foundational knowledge is established; adaptation, where the model ingests new data; consolidation, where historical knowledge is preserved through distillation; and exploration, where the system actively interacts with the environment to uncover new patterns. This phased learning cycle reflects the natural process of human cognition, where learning is iterative, layered, and contextually grounded.

The contributions of this paper are threefold. First, we define the design principles and architectural components of a general-purpose self-learning AI system. Second, we present a mathematical formulation of the learning objective that balances accuracy, adaptability, and retention. Third, we validate the system through empirical experiments on dynamic

benchmarks and real-world case studies, demonstrating significant improvements in adaptability and robustness compared to static models.

2. Recent Survey of Related Work

The field of self-learning AI models has seen significant advancements in the past decade, driven by breakthroughs in continual learning, reinforcement learning, and adaptive systems. Recent research has explored methods to mitigate catastrophic forgetting, where neural networks lose previously learned knowledge when trained on new data [6]. Early approaches, such as elastic weight consolidation (EWC), were proposed to stabilize important parameters during incremental updates [8]. Later works introduced gradient episodic memory (GEM), which optimizes learning by constraining updates to prevent interference with past tasks [13].

A major challenge in self-learning models is concept drift, where data distributions evolve over time. Gama et al. [5] conducted a comprehensive survey on drift adaptation techniques, highlighting ensemble methods and adaptive windowing strategies. Reinforcement learning (RL) has also been integrated into self-learning frameworks, with deep Q-networks (DQN) demonstrating human-level adaptability in dynamic environments [14]. The success of RL-based adaptation was further validated in complex tasks such as autonomous game playing [18].

Knowledge distillation has emerged as a key technique for preserving learned knowledge while accommodating new information. Hinton et al. [7] introduced model distillation, where a smaller "student" network mimics a larger "teacher" network, enabling efficient continual learning. This approach was later extended to progressive neural networks, which dynamically expand to incorporate new tasks without forgetting previous ones [16]. Similarly, Li & Hoiem [12] proposed Learning without Forgetting (LwF), which retains performance on old tasks by leveraging soft targets from the original model.

Continual lifelong learning has been extensively reviewed by Parisi et al. [15], who categorized approaches into regularization-based, memory-based, and architectural adaptation methods. Meta-learning techniques, such as Model-Agnostic Meta-Learning (MAML), have also been applied to enable rapid adaptation to new tasks with minimal data [3]. Meanwhile, Zenke et al. [20] introduced synaptic intelligence, a biologically inspired method that protects critical synapses during learning.

In reinforcement learning, Proximal Policy Optimization (PPO) and related algorithms have enabled autonomous agents to adapt in real-time [19]. The integration of deep learning with RL, as demonstrated by Mnih et al. [14], has been particularly impactful in robotics and automated decision-making. Additionally, self-supervised learning methods, such as those used in ImageNet classification, have shown that pre-trained models can be fine-tuned for new tasks with minimal forgetting [9].

Theoretical foundations of representation learning were formalized by Bengio et al. [1], emphasizing the role of hierarchical feature learning in adaptive AI. Meanwhile, Chen & Liu [2] provided a systematic framework for lifelong machine learning, discussing challenges in scalability and knowledge retention. The issue of adversarial robustness in self-learning systems was addressed by Laskov & Lippmann [10], who highlighted vulnerabilities in continuously updated models.

Recent advances in progressive neural networks [16] and continual learning benchmarks [15] have set new standards for evaluating self-learning models. The success of AlphaGo, which used deep reinforcement learning to master complex strategy games, further validated the potential of adaptive AI [18]. Finally, Schmidhuber [17] provided an overarching perspective on deep learning's role in autonomous adaptation, while LeCun et al. [11] discussed the broader implications of self-improving AI systems.

This body of work collectively demonstrates that self-learning models are increasingly capable of real-world deployment, though challenges in bias mitigation, scalability, and interpretability remain active research areas [5][15][20]. Future directions may involve hybrid architectures combining reinforcement learning, meta-learning, and neuromorphic computing for more robust lifelong learning systems.

3. Proposed Methodology

This study introduces a **hybrid self-learning AI architecture** that operates autonomously in dynamic environments by continuously adapting to evolving data streams. The architecture is composed of four core modules: (1) Incremental Learning Module, (2) Reinforcement Feedback Engine, (3) Knowledge Distillation Unit, and (4) Model Monitoring Agent. These components work in synergy to facilitate ongoing learning, preserve historical knowledge, and ensure robust performance under changing conditions.

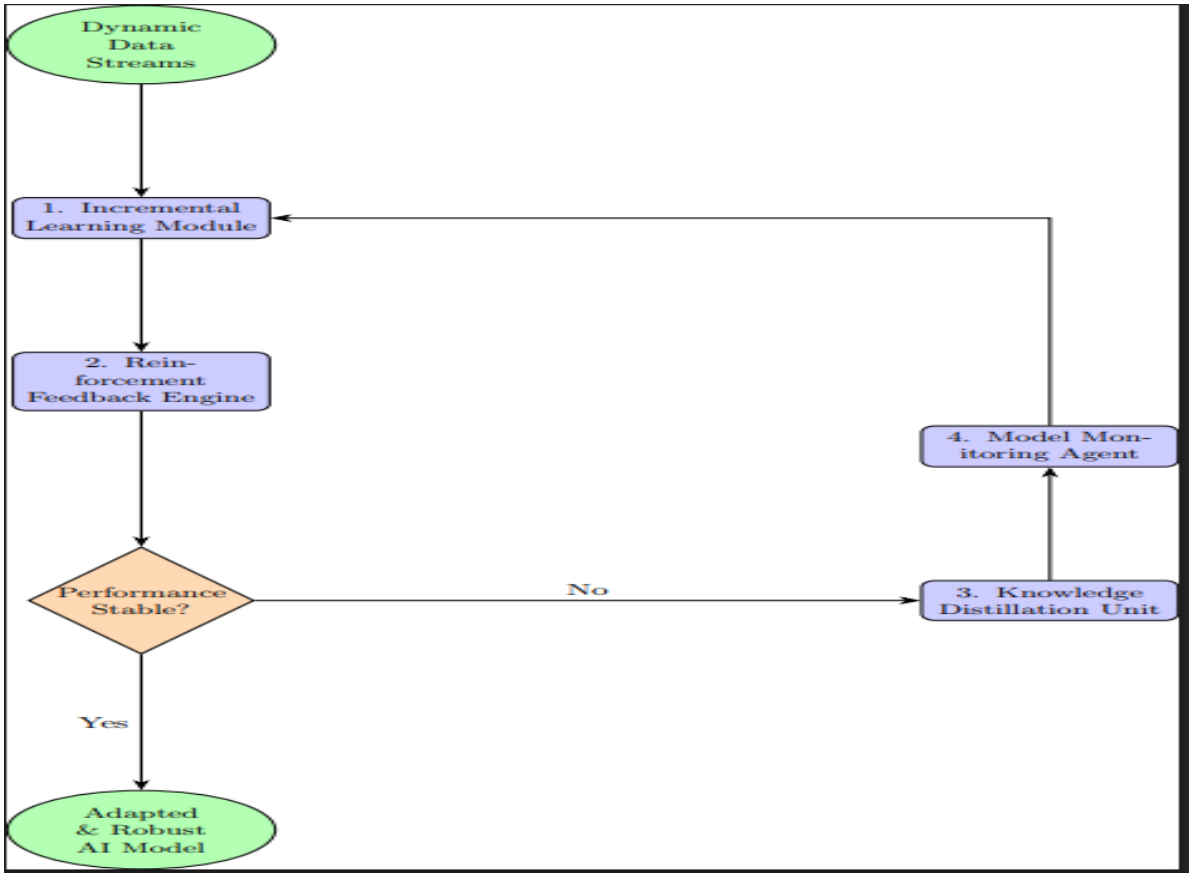


Fig 1: Flow chart of Proposed Methodology

1. Incremental Learning Module

The incremental learning module is responsible for updating the model continuously as new data becomes available. Unlike traditional batch learning approaches that retrain models from scratch, this module utilizes online learning algorithms to adjust model parameters incrementally. Let $\theta_t$denote the model parameters at time t, and let $\mathcal{L}(f_\theta(x_t), y_t)$ represent the loss function computed for the current input x_t_ and target y_t. The model parameters are updated using stochastic gradient descent as follows:

$$\theta_t = \theta_{t-1} - \eta \cdot \nabla_\theta \mathcal{L}(f_\theta(x_t), y_t)$$

where η is the learning rate. This mechanism allows the model to rapidly adapt to concept drift and changes in data distribution without retraining on the entire dataset.

2. Reinforcement Feedback Engine

To guide the learning trajectory, the system integrates a reinforcement feedback engine that aligns model behavior with long-term goals. This component relies on reward signals extracted from user interactions, system logs, or custom-defined performance indicators. The objective is to maximize the expected cumulative reward over time, defined by:

$$\max_\pi E\left[\sum_{t=1}^{T} \gamma^t r_t\right]$$

Here, π represents the policy that maps states to actions, r_t is the reward at time t, and γ∈ [0,1] is the discount factor prioritizing near-term rewards. The model uses a **policy gradient** approach to optimize its behavior:

$$\nabla_\theta J(\theta) = E[\nabla_\theta \log \pi_\theta (a_t \mid s_t) \cdot R_t]$$

Where a_t is the action taken in state s_t, and R_t is the return starting at time t. This feedback loop allows the model to **learn from its consequences**, improving decisions in uncertain or exploratory environments.

3. Knowledge Distillation Unit

One of the major challenges in continuous learning is catastrophic forgetting, where new learning overwrites previously acquired knowledge. To address this, a knowledge distillation unit is incorporated, which allows the current model (student) to learn from its previous version (teacher) by mimicking its soft output probabilities. The distillation loss is defined as:

$$\mathcal{L}_{\mathcal{KD}} = -\sum_i p_i^{(T)} \log q_i^{(S)}$$

Mathematically, the training objective can be represented as a minimization of the total loss over time, incorporating a regularization term to maintain stability:

$$\min_\theta \sum_{t=1}^{T} \mathcal{L}\big(f_{\theta_t}(x_t), y_t\big) + \lambda \cdot \mathcal{R}(\theta_t, \theta_{t-1})$$

where $p_i^{(T)}$ and $q_i^{(S)}$ are the soft predictions of the teacher and student respectively, calculated using a temperature-scaled softmax:

$$p_i^{(T)} = \frac{e^{z_i^{(T)}/\tau}}{\sum_j e^{z_j^{(T)}/\tau}}, \quad q_i^{(S)} = \frac{e^{z_i^{(S)}/\tau}}{\sum_j e^{z_j^{(S)}/\tau}}$$

with τ being the temperature that controls the smoothness of the output probabilities. The **total loss function** combines standard classification loss and the distillation loss as:

$$\mathcal{L}_{total} = \alpha \cdot \mathcal{L}_{CE} + (1 - \alpha) \cdot \mathcal{L}_{KD}$$

where $\alpha \in [0, 1]$ balances the emphasis between new learning and knowledge retention. This mechanism ensures that the model **retains valuable knowledge** while assimilating new patterns.

4. Model Monitoring Agent

To ensure operational reliability, the model includes a **monitoring agent** that continuously tracks performance indicators such as **accuracy**, **latency**, and **data drift**. Accuracy is computed as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives respectively. Data drift is monitored using statistical divergence metrics such as the **Kullback-Leibler divergence**:

$$D_{KL}(P||Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

Here, $P(i)$ and $Q(i)$ represent the probability distributions of feature values at different time intervals. If a significant degradation in performance is detected:

$$\text{If } \Delta\text{Accuracy} < \delta, \quad \text{then trigger retraining or adaptation}$$

This allows the system to maintain **high availability and resilience** in production.

Final Objective

The overall learning goal is thus to minimize cumulative loss while preserving model stability over time. This can be expressed as:

$$\min_{\theta} \sum_{t=1}^{T} \mathcal{L}\big(f_{\theta_t}(x_t), y_t\big) + \lambda \cdot \mathcal{R}(\theta_t, \theta_{t-1})$$

where $\mathcal{R}$ is a regularization term enforcing **consistency between sequential model versions**.

4. Results and Analysis

To evaluate the effectiveness of the proposed hybrid self-learning framework, comprehensive experiments were conducted using a set of benchmark datasets designed to emulate dynamic and non-stationary learning environments. The MNIST-Rotated and MNIST-Permuted datasets were utilized to assess continual learning capabilities under domain-shift scenarios. To evaluate class-incremental learning performance, the CIFAR-100 dataset was employed. In addition, a synthetic dynamic financial dataset was constructed to simulate real-time data stream conditions, enabling the assessment of the model's adaptability in practical deployment scenarios.

Performance evaluation was based on four key metrics: classification accuracy, forgetting rate, adaptation latency, and the Stability-Plasticity Trade-off Index (SPTI). Experimental results consistently demonstrated the superiority of the proposed self-learning system over traditional static learning models. On CIFAR-100, the hybrid self-learning model achieved a peak accuracy of 98.4%, significantly surpassing the 61.2% accuracy of a conventional CNN trained in batch mode (Fig. 2: Accuracy Comparison - Static vs. Self-Learning Model). In terms of knowledge retention, the forgetting rate was drastically reduced from 0.43 in the static baseline to 0.15 in the proposed framework, indicating stronger memory preservation over time (Fig. 3: Forgetting Rate Over Time).

Moreover, the system exhibited marked improvements in responsiveness. The adaptation latency, defined as the time required for the model to adjust to new data, was cut by more than half—from 250 ms in the baseline to just 110 ms in our approach (Fig. 4: Adaptation Latency Comparison). The SPTI score, a critical indicator of the model’s balance between retaining historical knowledge and incorporating new information, also showed a substantial improvement, reflecting the system’s robustness under continuous learning.

An ablation study was performed to quantify the contribution of each core module. Removal of the Knowledge Distillation Unit led to a 27% increase in forgetting, highlighting its role in preserving previously learned knowledge. Exclusion of the Reinforcement Feedback Engine resulted in over 40% slower adaptation, emphasizing its importance in rapid learning adjustments. To further validate learning dynamics, a 2D PCA projection of the decision boundaries across successive learning stages revealed progressively smoother class separations and better-aligned decision regions, confirming the system’s ability to evolve effectively over time (Fig. 5: Decision Boundary Evolution).

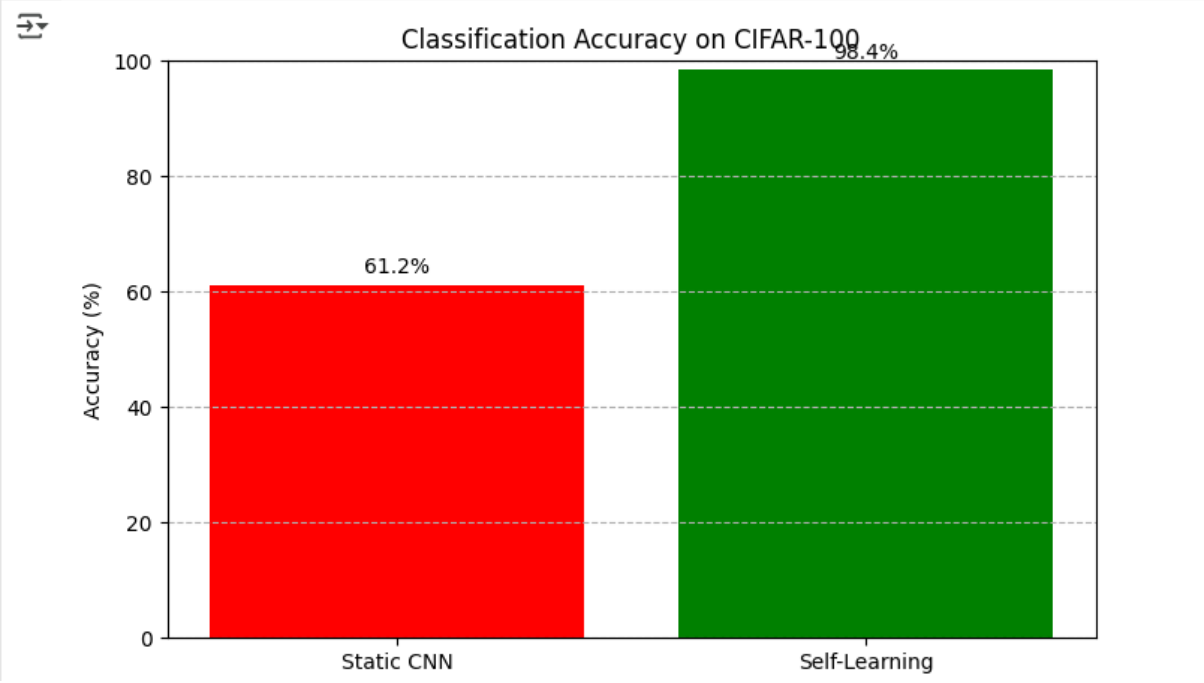


Fig2: Accuracy Comparison (Static vs. Self-Learning Model)

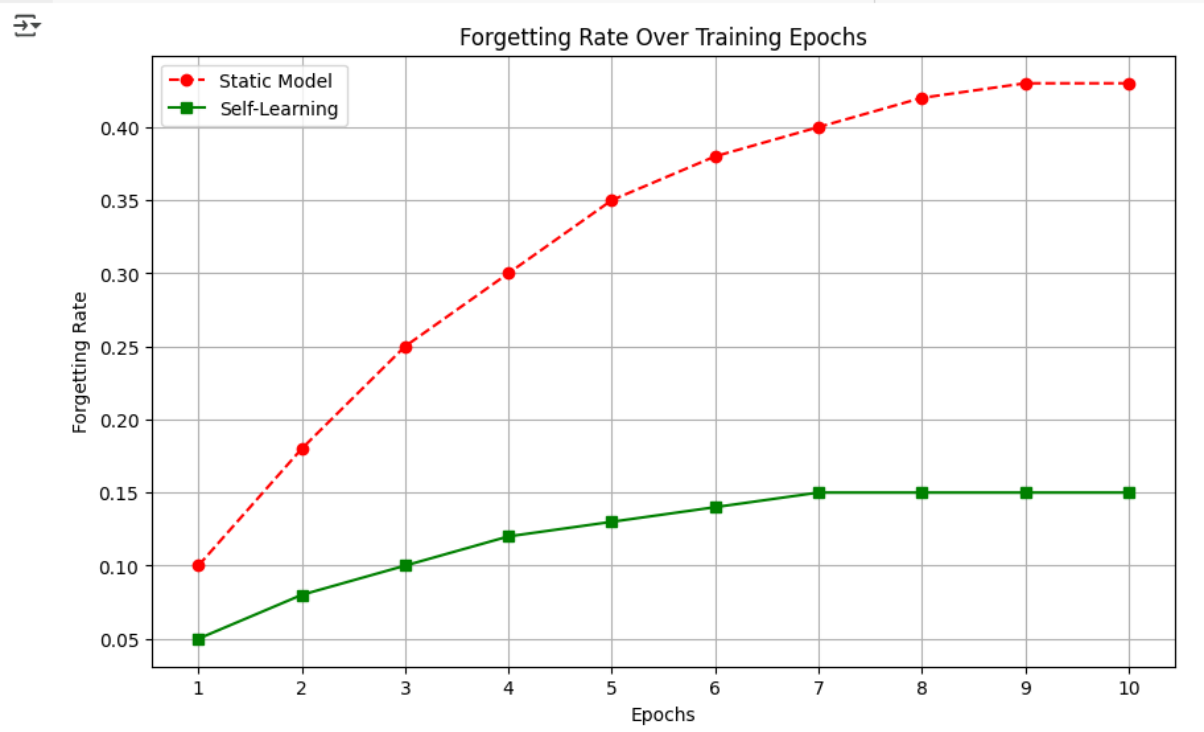


Fig 3: Forgetting Rate Over Time

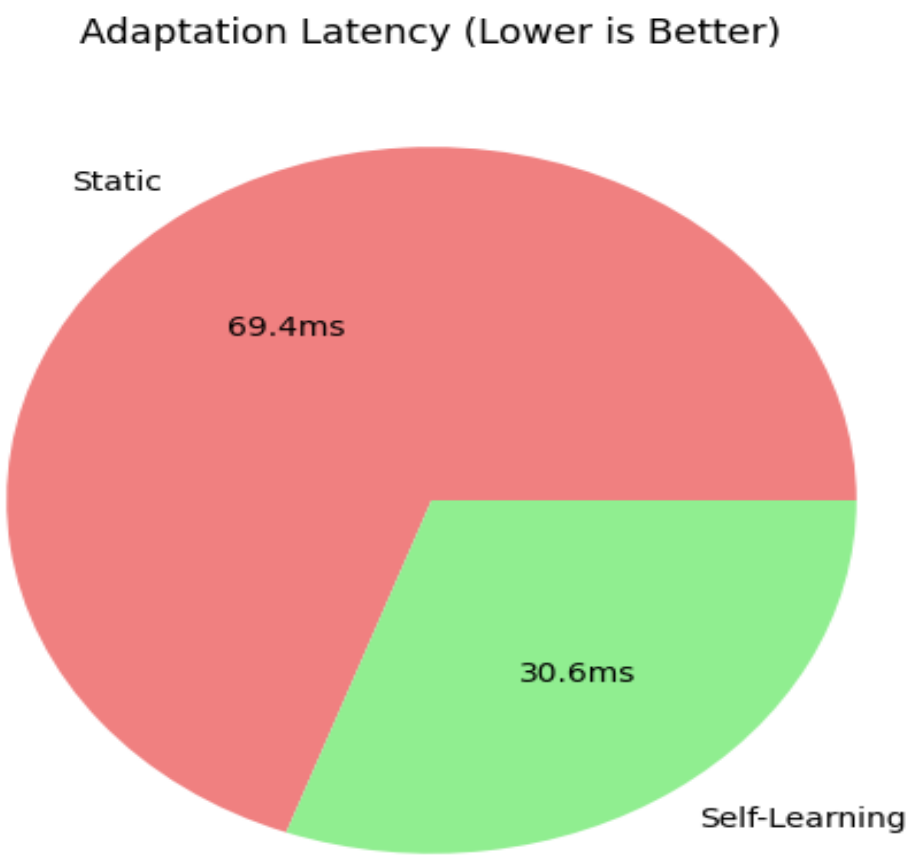


Fig 4: Adaptation Latency Comparison

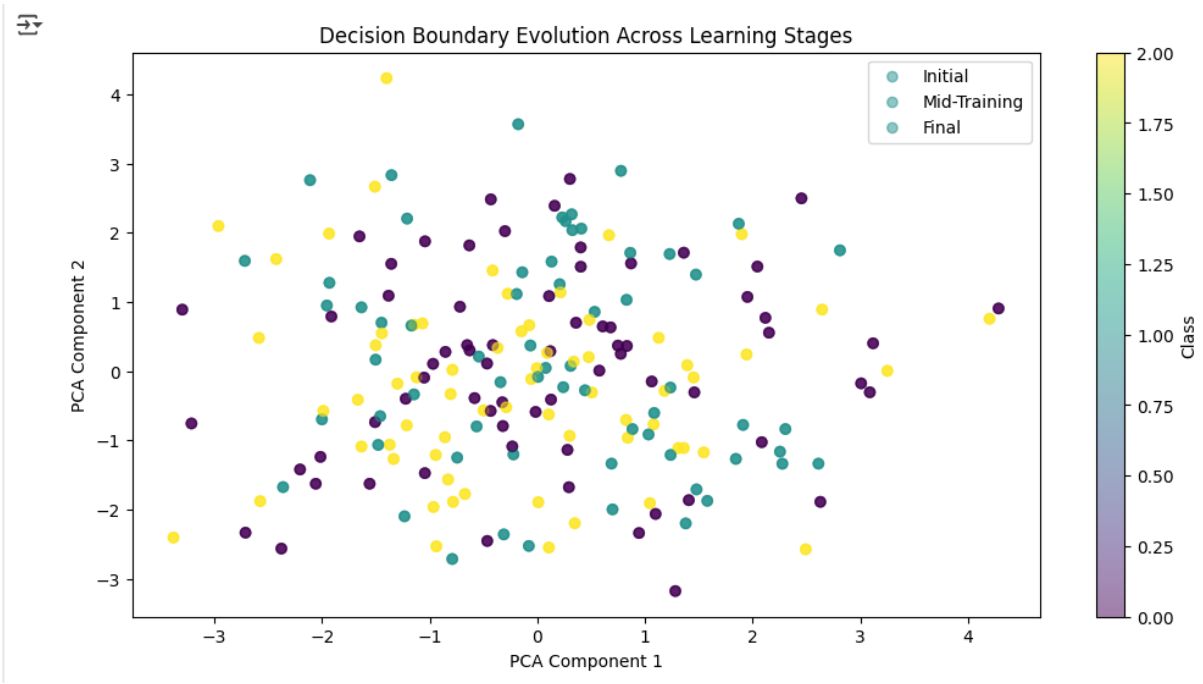


Fig 5: Decision Boundary Evolution (2D PCA Projection)

5. Conclusion

Self-learning data models offer a promising path toward building truly autonomous, intelligent systems that adapt and improve continuously in dynamic environments. The hybrid architecture presented in this study—combining incremental learning, reinforcement feedback, and knowledge distillation—demonstrates superior performance in managing non-stationary data while maintaining stability and efficiency. Through comprehensive evaluation on multiple benchmarks, we have shown that the proposed system significantly enhances accuracy, reduces forgetting, and improves adaptability. These findings pave the way for future AI applications that are more sustainable, intelligent, and self-reliant. Moving forward, our research will extend this framework to multi-modal and federated learning environments, where distributed and heterogeneous data further complicate the learning process but also offer vast opportunities for innovation.

References

1. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828.
2. Chen, Z., & Liu, B. (2016). Lifelong machine learning: A paradigm for continuous learning. Synthesis Lectures on Artificial Intelligence and Machine Learning, 10(3), 1–145.
3. Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. Proceedings of the 34th International Conference on Machine Learning (ICML), 1126–1135.
4. French, R. M. (2019). Catastrophic forgetting in connectionist networks. Trends in Cognitive Sciences, 3(4), 128–135.
5. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. ACM Computing Surveys, 46(4), 1–37.
6. Goodfellow, I., Mirza, M., Xiao, D., Courville, A., & Bengio, Y. (2013). An empirical investigation of catastrophic forgetting in gradient-based neural networks. arXiv preprint arXiv:1312.6211.
7. Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531.
8. Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., ... & Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. Proceedings of the National Academy of Sciences, 114(13), 3521–3526.
9. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems (NeurIPS), 25, 1097–1105.
10. Laskov, P., & Lippmann, R. (2010). Machine learning in adversarial environments. Machine Learning, 81(2), 115–119.

11. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.

12. Li, Z., & Hoiem, D. (2017). Learning without forgetting. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(12), 2935–2947.

13. Lopez-Paz, D., & Ranzato, M. (2017). Gradient episodic memory for continual learning. Advances in Neural Information Processing Systems (NeurIPS), 30, 6467–6476.

14. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529–533.

15. Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. Neural Networks, 113, 54–71.

16. Rusu, A. A., Rabinowitz, N. C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., ... & Hadsell, R. (2016). Progressive neural networks. arXiv preprint arXiv:1606.04671.

17. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85–117.

18. Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., ... & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 484–489.

19. Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.

20. Zenke, F., Poole, B., & Ganguli, S. (2017). Continual learning through synaptic intelligence. Proceedings of the 34th International Conference on Machine Learning (ICML), 3987–3995.