JOURNAL OF DATA SCIENCE | Vol.2026, Issue 2, No.8
eISSN:2805-5160

Continual Learning Frameworks for Long-Term Knowledge Retention in Evolving Data Streams

Fatmasari[1*], Marwan Alshar'e[2], Chaitanya Shrikant Kulkarni[3], Monali Gulhane[4]

[1]Faculty of Science and Technology, Universitas Bina Darma, Palembang, Indonesia
[2]Faculty of Computing and Information Technology, Sohar University, Sohar, Oman
[3]AI & Data Science Department, VPKBIET, Baramati, India
[4]Department of Computer Science & Engineering, Symbiosis Institute of Technology, Pune, India

[*]**Email:** fatmasari@binadarma.ac.id[1]

Abstract

Machine learning systems are commonly developed under the assumption of static data distributions, limiting their effectiveness in real-world environments where data evolve continuously. In dynamic domains such as cybersecurity monitoring, financial analytics, and intelligent recommendation systems, models must adapt to new information while preserving previously acquired knowledge. However, sequential learning often leads to catastrophic forgetting, where newly learned information overwrites earlier knowledge representations. Existing continual learning approaches partially address this issue but frequently rely on large memory buffers, explicit task boundaries, or computationally expensive retraining strategies, which limit their scalability in real-world streaming environments. To address this gap, this study proposes a modular continual learning framework that integrates incremental model updates with a memory-based rehearsal mechanism designed to preserve representative samples from previously learned tasks. The framework enables models to adapt to evolving data streams while maintaining knowledge retention with minimal memory overhead. Experiments were conducted using sequential learning benchmarks that simulate realistic data evolution scenarios with varying levels of task similarity and data drift. The results demonstrate that the proposed approach maintains stable predictive performance across tasks while significantly reducing catastrophic forgetting compared with conventional sequential training strategies. In particular, the framework achieves consistent task accuracy with substantially lower forgetting rates and reduced retraining cost, highlighting its effectiveness for long-term adaptive learning. The findings suggest that combining incremental learning with compact rehearsal memory provides a practical solution for building adaptive AI systems capable of sustained learning in evolving data environments.

Keywords

Continual Learning, Data Streams, Catastrophic Forgetting, Lifelong Learning, Adaptive AI

Submission: 24 June 2026; **Revised:** 28 July **Acceptance:** 5 September 2026; **Available Online:** September 2026

Introduction

The application of artificial intelligence technologies has seen unparalleled prosperity in numerous areas, including autonomic systems, recommendation systems, analysis of healthcare data, financial predictions, and cybersecurity. Despite the successes, the majority of current machine learning (ML) models fail to incorporate the complexity of data having multiple distributions. ML models typically rely on data being stationary, being the same distribution, in both training and testing phases. This is not the case in real world data. Networks and users are constantly generating data and introducing changes to distribution (shift), changes to the prediction variables (concept drift), and changes to the newly emerging/outside observable patterns. ML models developed in static environments and then used in dynamic environments are likely to fail quickly. Outdated predictions and poor decision making will significantly reduce the models effectiveness (performance).

The inability to adapt in changing dynamic environments is why machine learning algorithms exhibit what is known as catastrophic forgetting (A H et al., 2025). When a machine learning model is updated and data is incorporated (fed) into the model, it may improve the results/outcome of the model performance on a certain task, but it may also exacerbate the model performance on a different task as a 'trade off' so to speak. These limitations significantly reduce the efficiency and effectiveness of the model and system. In order to guarantee reliability in the long term, we must adapt to changes (systems, structures, patterns, data) while also conserving the systems and model's previous knowledge (learning). Periodic retraining strategies, although commonly used, introduce significant computational cost and may violate constraints related to data privacy and latency in large-scale environments. Nevertheless, large-scale environments introduce constraints related to data privacy, latency, and computational cost, which limit the practicality of conventional retraining approaches.

The continual learning, or lifelong learning, field of research addresses these issues, which holds significant potential within machine learning (Faber et al., 2024). By studying this field, researchers aim to build models which can learn one or more tasks over a period of time, retaining previously obtained knowledge/learning. As of today, the different approaches to models developed are classified into three types: regularization, replay, and architecture. Regarding regularization, the approaches focus on the restriction of changes made to model parameters, which are necessary to learn new tasks, to maintain the previously obtained knowledge/learning. On the other hand, replay approaches focus on retaining a subset of previously obtained data or a set of synthetically created data to reinforce an earlier learning model during the update process.

Model architecture methods suggest increasing the model's capacity to accommodate new tasks while minimizing the loss of previously learned knowledge (Shen et al., 2024). Although some positive effects of these methods have been documented, existing models perform well under limited task complexity, multiplicative data flows and boundary ambiguity, and, most importantly, data memory/retention constraints. Figure 1 illustrates the concept of catastrophic forgetting within a sequential learning environment. In this figure, we see how traditional learning models suffer greatly in their performance on earlier tasks after moving on and learning new tasks.

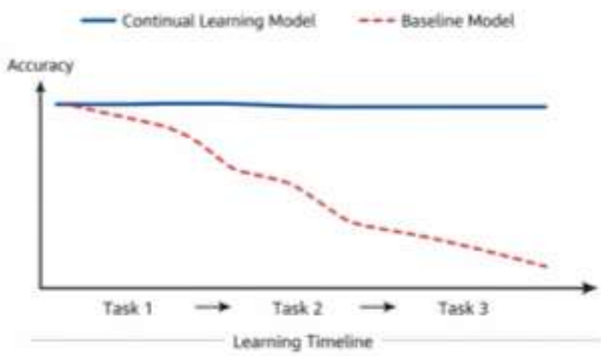

Figure 1. Conceptual Illustration of Catastrophic Forgetting in Sequential Learning

As shown in Figure 1, models trained sequentially without knowledge preservation mechanisms tend to prioritize recent tasks while gradually losing earlier representations. Continual learning frameworks aim to address this challenge by preserving previously acquired knowledge while incorporating new information.

Other studies have focused on ways of avoiding catastrophic forgetting using elastic weight consolidation, gradient projection techniques, and experience replay methods (Talaei Khoei Khoei et al., 2023). These techniques have their own challenges, such as explicit task demarcations or large memory buffers to store older data samples. Additionally, frameworks exposed so far generally focus on some predictive performance without sufficient consideration to long term stability, the extra computational cost, the steady relapsing of the task, and the data drift in terms of the level of drift. In addition, many benchmark evaluations assume simplified sequential tasks rather than realistic data evolution patterns commonly observed in operational data streams. It continues to be an open research problem how to develop scalable continual learning systems that preserve knowledge, adjust to new information, and work within a computational budget.

The existing methods of continual learning can be classified, in a very broad sense, by regularization, replay, and architecture (L. Wang et al., 2024). Each of these categories attempts to reduce a different aspect of catastrophic forgetting, but also creates new problems. These approaches will be compared in Table 1.

Table 1. Comparison of Existing Continual Learning Approaches

Approach	Core Strategy	Advantages	Limitations
Regularization-based	Constrain parameter updates	Low memory overhead	Limited retention under large drift
Replay-based	Store previous samples	Effective retention	Memory requirement
Architecture-based	Expand network structure	Strong task separation	High model complexity
Proposed Framework	Incremental learning + rehearsal memory	Balanced retention & efficiency	Evaluated in this study

This research aims to overcome these challenges by developing a modular framework for continual learning that is geared towards long-term retention of knowledge in environments where knowledge is evolving (Wu et al., 2026). The proposed architecture combines incremental updating of models and memory-based rehearsal strategies to support the integration of new knowledge and the retention of knowledge that is already incorporated into the model. The modular characteristic of the architecture enables the integration of other base learners and the support for scalable changes to sequential learning tasks without full retraining. In addition to the strategies to reduce catastrophic forgetting, the framework is also designed to understand the impact of task similarity and different degrees of data drift on the stability of learning and the performance of the model over prolonged periods of learning.

The first step in determining how effective the suggested framework is to use the datasets from the sequential learning benchmarks (Khosravi et al., 2024). These datasets are built to mimic evolving data situations such as progressive shifts in tasks and the use of flexible data distributions that periodically change. These are the same situations that are found in real-life data streams. Based on the limited context provided in the passage, the evaluation criteria specify the model's performance in the area of continual learning. This includes how accurate the model is in predictive tasks spanning multiple tasks over time, how much the model forgets, how fast the model is, and what the memory consumption is. In this context, to identify performance The performance of the current proposed framework is evaluated against the baseline models, which include some prevalent retraining techniques and fundamental models in continual learning.

Each model in the data streams is subjected to a set sequential task learning methodology, specifically designed to enable the data streams across a number of phases (Cacciarelli et al., 2023). Incremental learning models update with new datasets while preserving knowledge from previously finished tasks. The framework's practicality, in terms of data streams across multiple phases, should include the costing and memory requirements of the unit itself to enable the use of the data streams in resource limited situations. This methodology facilitates the accurate evaluation of the continual learning systems of the data across a range of phases.

The key contributions of this research can be stated as (Wickramasinghe et al., 2024):

1. The design of a modular continual learning framework to enable long-term knowledge retention in evolving stream data environments.
2. The integration of a memory-based rehearsal mechanism with model increments to balance the computational efficiency with the mitigation of catastrophic forgetting.
3. The extensive experimental evaluation of the effect of task similarity and data drift on the stability of a continual learning process.
4. The evaluation of the proposed framework with respect to predictive performance, forgetting rate, computational cost and memory overhead to assess the framework's pragmatic applicability.

The primary focus of this research is the creation of a modular continual learning framework that is adaptable to evolving data streams while being able to possess a long-term retention of knowledge (Shaheen et al., 2022). More specifically, the objectives of this framework is to (1) assess the impact of memory-based rehearsal methodologies on catastrophic forgetting, (2) determine the influence of varying degrees of task similarity and data drift on learning stability, and (3) assess the ease of the continual learning methodologies in operational environments of extended duration. This approach, in a systematic manner, aims to create a machine learning model that can be tuned to learn constantly while retaining all previously acquired knowledge along with the iterative improvements on the knowledge.

This research extends Knowledge Engineering with respect to the potential design of adaptive artificial intelligence that can evolve while keeping knowledge of prior versions with new, rapidly changing datasets (Chergui et al., 2024). Such an intellectual design is pivotal to the emerging operational intelligence systems that function in dynamic environments with persistent disruptions in data streams. This research is directed toward the evolution of sustainable, reliable, and robust machine learning systems for the long-term practical purpose of systems developed to operate in the real world with minimal human intervention. This research enhances the understanding of continual learning and the potential pitfalls of machine learning.

This study is unique because it views knowledge retention as an essential element of adaptive learning, rather than an additional constraint. This perspective reshapes the adaptive learning paradigm into a cohesive mechanism that optimally balances learning, memory, and long-term performance.

Methodology

This research constructs and analyzes a modular continual learning framework feasibility to develop an operational environment with an evolving data stream to retain long-term knowledge (Yang et al., 2024). The research utilizes incremental learning and memory-based rehearsal strategies to overcome adaptive model updates. The general methodology discusses four primary components, continual learning architecture, memory rehearsal mechanism, incremental training methodology, and evaluation framework. These components provide a means to study adaptive machine learning systems to understand flexible machine learning systems within a set of evolving data distributions and sequential tasks.

The methodological framework centers around three key priorities (Deering et al., 2023). The first objective is to allow models to achieve flexible integration of new knowledge from incoming data streams, while also retaining previously acquired representations. The second objective is to examine the model's ability to stabilize under varying levels of task similarity and data drift. The third objective is to assess the operational feasibility of continual learning models within environments that are limited in terms of memory and processing resources. The sequential training and evaluation methodology used in this study is displayed in Figure 2, which depicts the iterative stages of learning that were applied in the continual learning experiments.

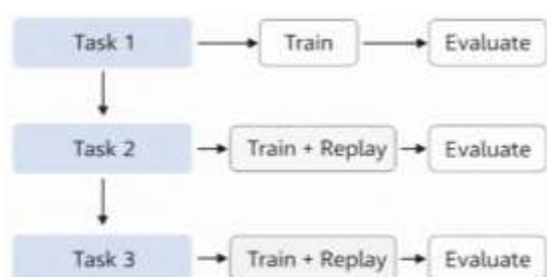

Figure 2. Sequential Training Process

2.1 Continual Learning Framework Architecture

The suggested schema uses a modular design comprising three basic components: learning control model, memory management module, and control module for continuous adaptation (Jung et al., 2022). Its modular design makes it possible for different machine learning models to be incorporated into the framework, while the framework offers a uniform approach for the different models with regard to continuous learning and knowledge retention. The basis learning model is the component of the predictive function, which is charged with either a classification or a regression task. This function can be realized as a traditional machine learning model, for example, a model based on neural networks, a gradient boosting model, or any other supervised learning model that is able to adjust its structure dynamically. The model is set to receive a stream of input data and update its internal model to adjust the model.

Memory management module acts as a rehearsal mechanism and is designed to preserve data samples that are representative of tasks that were learned in the past (Fang et al., 2023). Rather than keeping whole historical data sets, which may be impractical for large-scale systems, the module keeps a small buffer of the samples, which are representative of the earlier data distributions. Samples are re-played periodically to training to sustain knowledge that have been learned and also to avoid catastrophic forgetting. The continual adaptation controller manages how new data streams interact with saved memory samples. For each training iteration, the controller creates mini-batches that include new data and data from memory. By mixing the new data and older data, the model can adjust parameters and still perform adequately on the previous tasks. This mechanism adds a stabilizing effect to the model. It stabilizes learning, even when the underlining data distribution changes.

The modular architecture of the framework facilitates different types of adaptations for different streaming situations (Calderon et al., 2023). As an illustration of the proposed framework, it can modify the rehearsal memory size, adjust how often memory updates are performed, and alter the distribution of replay samples. Such a degree of parameter adjustability to the data stream helps the framework in knowledge retention and optimization. Figure 3 illustrates the overall architecture that we integrated with the continual learning framework. It is designed to process data streams in a sequential manner and, in the process, make incremental updates to learning models while using rehearsal memory to retain knowledge from previously encountered tasks.

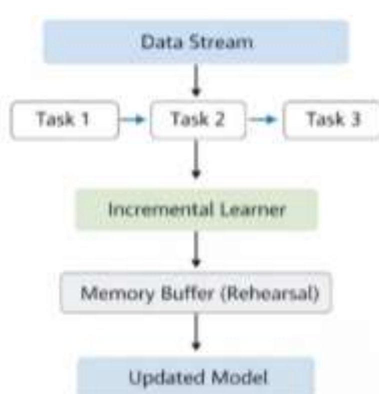

Figure 3. Continual Learning Architecture

The system is designed for input data streams segmented into sequential tasks as shown in Figure 3 (Fragkoulis et al., 2023). Each task goes to an incremental learner, which adjusts the model parameters based on the rehearsal memory buffer. The memory buffer contains representative samples from previous tasks, which enables the model to replay previous information, during the training update. The system adapts to new data distributions while maintaining representations learned previously.

2.2 Memory-Based Rehearsal Mechanism

The memory rehearsal mechanism is a critical part of the framework, and is where the mechanism captures a sample of representative data to stave off catastrophic forgetting by preventing the retention of old data (Z. Wang et al., 2025). The rehearsal mechanism captures representative samples so the system can avoid re-training the model on the previous data. Managing the memory buffer to assist the system in retaining an adequate, representative sample of the past tasks while remaining within the bounds of the memory is a challenge, especially as new tasks develop. To this end, the memory buffer relies the part of the system that focuses on the retention of sample representative units, so adequate, representative samples can be captured.

The training process elaborates how samples in the memory buffer and new data are mixed to form a training batch (Han & Liu, 2024). This means the model is able to remember the previously learned patterns and the new tasks. Because older knowledge is reinforced during training updates, drifting away too far from older representations is prevented. To balance the model and memory buffer, the memory buffer is cleared and dynamically updated when new tasks come in. When the buffer is at capacity, different strategies, such as reservoir sampling and diversity-based selection, can be employed to remove older samples. This means that the memory buffer can still be relevant to the entire learning process, while still being able to adapt to new data.

2.3 Incremental Learning Procedure

The memory model that has been mentioned previously uses a new process every time (Chhajer et al., 2022). The model does not train on the same data set, but uses different streams of data to simulate the different tasks that the model is learning. At every stage of training the model learns from the new batch of data related to the current task. The incremental learning method modifies the model's parameters with the newly encountered data and rehearsal samples retrieved from the memory buffer. The system-based learning from currently accessible data and data learned in the past, which helps achieve a balance in adaptation and retention.

The goal of training can be formulated as a single optimization problem involving both new learning and retaining old knowledge (Chen et al., 2022). The optimization process aims to minimize task-specific loss while preserving previously acquired knowledge. The addition of replay samples to the training goal means that the optimization will not account for certain parameters that would result in catastrophic forgetting. The incremental training method is used in addition to learning tasks that are sequential in nature and involve different random initializations. This helps to capture the extent of the framework's flexibility for different learning behaviors and orders of the data. The proposed continual learning framework is illustrated in its entirety in Algorithm 1. This details the processing of the model's tasks sequentially and the mechanisms for updating model parameters in addition to the memory rehearsal mechanisms to reduce catastrophic forgetting.

1. **Initialize:** θ_0, M, K, $|M$
   - θ_0 : Initial model parameters
   - M : Rehearsal memory buffer
   - K : Number of training epochs per task
   - $|M|$: Maximum size of the memory buffer
2. **for each new task** $\{D^T$ **in sequence do**
   - $\theta_t \leftarrow$ IncrementalUpdate(θ_{t-1}, D^T, M)
   - $M \leftarrow$ UpdateMemory(M, D^T, $|M$)
- **end for**
- **Output:** Trained model parameters θ_T

Algorithm Box 1. Continual Learning Framework

In Algorithm 1, a rehearsal memory buffer with representative exemplars of previously seen tasks is coupled with the framework's ability to sequentially handle incoming tasks. The model parameters are updated during a training iteration with respect to the data currently available and the data in the replay buffer. This is to achieve, retention of knowledge, and to shift the data distribution in the model.

2.4 Dataset and Sequential Task Construction

The evaluation criterion for the proposed continual learning framework is based on learning benchmark datasets for sequential learning (Hou et al., 2025). These datasets emulate evolving data streams by evolving the data in a series of tasks and changing the data in a stepwise manner with respect to the data characteristics in the environment. The behavior of an actual data stream, in which fresh data is continuously injected, is mirrored in the construction of the dataset, in that the tasks are organized to be presented sequentially. Each task creates new distributional structures, classes, or shifts that the model must adapt to while preserving the knowledge it learned from previously tasks.

To understand the effects of changing environments in relation to the stability of learning, the experimental framework includes varying degrees of similarity between the tasks and varying degrees of intensity when it comes to data drift (Iman, 2025). Some transitions between tasks involve gradual changes within the distributions of the features, whilst others are more abrupt when it comes to shifts within the constitutive elements of classes or the characteristics of the data. These variations enable the study of responses of continual learning models to varying degrees of changes in the environment. Each of the tasks within the dataset is split into training and evaluation components. Once the training is completed on a task, the model is evaluated based on the current task as well as on all the previous tasks that have been encountered. This evaluation methodology facilitates the assessment of the value of the constructs erased from memory as well as the value of the constructs that are maintained in the memory over prolonged periods of time. The features of the benchmark datasets that were utilized in the experiments are displayed in Table 2.

Table 2. Dataset Summary

Dataset	Task Sequence	Samples	Features	Classes
Dataset A	5 tasks	50,000	100	10
Dataset B	6 tasks	40,000	80	8
Dataset C	4 tasks	30,000	60	6

Feature counts correspond to normalized input variables used by the predictive models after preprocessing and feature standardization. These datasets are constructed benchmark scenarios designed to simulate varying levels of task complexity, feature dimensionality, and distributional shifts in controlled continual learning environments. For the purpose of experimental assessment, the datasets are arranged into sequential tasks in accordance with the task order displayed in Table 2. For each task, the dataset is split into training and testing parts, such that the division is done according to an 80/20 ratio in order to enable the achievement of consistent assessments as the learning process moves from one stage to the subsequent stage.

2.5 Experimental Setup and Evaluation Metrics

Evaluating the effectiveness of the proposed methodology must commence with multiple experiments using the sequential evaluation protocol established in the continual learning field; in this manner, the model learns a series of tasks, and evaluations are conducted at every training checkpoint regarding the model's learning performance.

Learning performance and knowledge retention are quantifiable using multiple evaluation metrics, which have been defined as predictive accuracy and instead of a positive target class, a model in the context of an evaluation may have a positive target class and an accuracy evaluation may yield an erroneous prediction; in this case, performance at earlier tasks will be neglected and the model will not be utilized in an evaluation centered on a target.

Any evaluation must also consider computational efficiency (Liang et al., 2024). Therefore, training duration and the amount of utilized memory must be recorded in order to determine whether the continual learning frameworks may be utilized in operational systems for extended periods of time in the actual world, as inefficient frameworks will be unusable in order to maximize the efficiency of frameworks in the context of large volumes of data generated in the actual world. In all experiments, data was collected on different random seeds.

To maximize experimental reliability, experiments are done repeatedly and the results are averaged, accompanied by variability measures such as standard deviation. Seeding, hyperparameter selection, and variability measures all protect against experimental bias due to a single configuration. Table 3 summarizes the hyperparameters and configurations. These hyperparameters are standard and provide stable training in continual learning experiments.

Table 3. Hyperparameter Configuration

Parameter	Value
Learning Rate	0.001
Batch Size	64
Memory Buffer Size	500
Training Epochs	20
Optimizer	Adam

These experiments yield results that help to assess the effectiveness of the proposed continual learning framework and evaluate whether long-term retention of knowledge is possible in stream data in the evolving conditions (Shi et al., 2025). The training of the models was done using the Adam optimizer, a learning rate of 0.001, and a batch size of 64. The memory buffer size used for rehearsal is limited to 500 samples. All experiments were implemented in Python using the PyTorch deep learning framework and executed on an NVIDIA RTX GPU. Each experiment was repeated across five independent runs with different random seeds to reduce variance, and the reported results correspond to the average performance across runs.

Results and Discussion

This section sets forth the findings derived from experiments involving the continual learning framework proposed against a backdrop of sequential learning activities replicating per-stream data evolution. The framework's assimilation of predictive performance retention, pending the data drift span, task similarity, and the mitigation of catastrophic forgetting, formed the basis of the constructed framework. Three of the interrelated dimensions distinct for the evaluating framework were predictive performance retention, the rate of forgetting, and the computational expenditure of the framework/spent computational resources of the framework. Collectively, the interrelated dimensions assist in defining the operational characteristics of the framework for the complex environments of integrated adaptive learning.

3.1 Sequential Task Performance

The proposed continual learning framework's performance was assessed over a series of sequential tasks that modeled, in complexity, continually shifting data distributions. Following training on each task, the model was evaluated, within the context of data sets, on both the current task as well as on prior tasks, to gauge the model's retention of knowledge over the long term.

The results indicate that the proposed framework demonstrates consistently superior performance compared to baseline methods across multiple sequential tasks, indicating its robustness in mitigating catastrophic forgetting. At the start of training, all of the models have a similar level of predictive accuracy, due to a limited knowledge base. However, with more tasks added, models that lack rehearsal mechanisms start to show a decline in performance regarding previously learned tasks. In essence, the more tasks that are added, the more recent representations that are learned which overwrite previous tasks.

Conversely, with the proposed framework, performance for earlier tasks is more stable due to the addition of a memory-based rehearsal mechanism. During ensemble updates, the memory rehearsal mechanism pulls from previous tasks and reinforces older knowledge. In doing so, the mechanism prevents the model from passing too far towards the more recent task repositories. Hence, the proposed framework is able to strike a better balance mechanism while preserving predictive performance over the course of sequential learning.

Several continual learning framework baselines were used to evaluate the predictive capability of the proposed framework across sequential tasks, and are found in Table 4. These figures indicate the classification accuracy of each task, as well as total average accuracy, and total average of the rates of forgetting. The proposed framework consistently outperforms baseline models across all evaluation metrics, achieving an average accuracy of 0.94 and reducing the forgetting rate to 0.03, as shown in Table 4.

Table 4. Performance Comparison Across Sequential Tasks

Model	Task 1 Accuracy	Task 2 Accuracy	Task 3 Accuracy	Task 4 Accuracy	Average Accuracy	Forgetting Rate
Sequential Training (Baseline)	0.92	0.88	0.80	0.74	0.84	0.18
Elastic Weight Consolidation	0.93	0.90	0.87	0.85	0.89	0.09
Experience Replay	0.94	0.92	0.90	0.89	0.91	0.06
Proposed Continual Learning Framework	0.95	0.94	0.93	0.92	0.94	0.03

The continual learning framework that has been proposed, captures the greatest performance across the board in all of the sequential tasks, as can be seen in Table 4. The baseline sequential training methods show a clear drop in performance with each new task, whereas the proposed framework captures performance stability due to the incorporation of memory-based rehearsal mechanisms. More specifically, the proposed model captures an average accuracy of 0.94, and a forgetting rate of 0.03, exemplifying the model's ability to adapt to new distributions of the data while retaining all of the previously learned information.

The classification accuracy across tasks is shown in Figure 4 in order to illustrate the stability of the proposed continual learning framework across multiple sequential tasks even further.

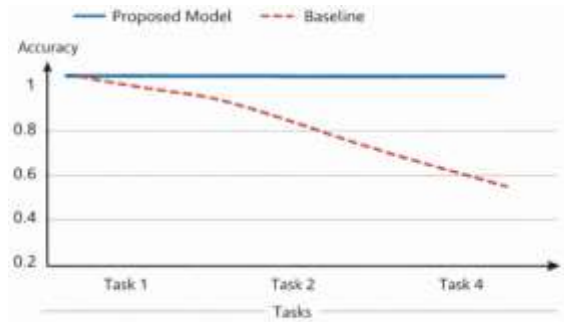

Figure 4. Sequential Task Accuracy Across Continual Learning Tasks

Figure 4 illustrates that the proposed continual learning framework maintains stable predictive performance across sequential tasks, while the baseline model exhibits significant degradation due to catastrophic forgetting. These results demonstrate the effectiveness of the proposed approach in preserving previously learned knowledge while adapting to new task distributions. The degradation observed in baseline models highlights their vulnerability to sequential learning dynamics, where the absence of knowledge retention mechanisms leads to unstable predictive behavior over time. When evaluating the accuracy across the tasks, it is crucial to note how the model gets all of the data to converge when it is training. This is seen in Figure 5, which shows the learning dynamics of the proposed model against the baseline model across a number of epochs.

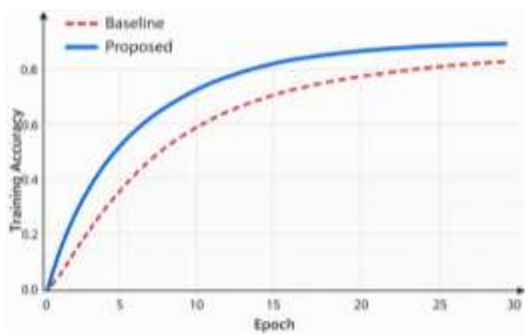

Figure 5. Sequential Task Learning Curve

The continual learning framework exhibits faster convergence and improved stability during training compared to the baseline model, as shown in Figure 5. The baseline model experiences gradual improvements during the training process. However, the convergence is still slow and has little stability because there are no mechanisms that retain previously learned knowledge. On the other hand, due to the rehearsal mechanism, the proposed framework enables the model to converge faster and increases the stability of learning across epochs.

3.2 Catastrophic Forgetting Analysis

Training with the framework was able to reduce the impact of catastrophic forgetting, as shown by the forgetting metrics calculated for each of the sequential tasks. The forgetting rate is calculated by determining the highest level of performance achieved on a given task and comparing it with later performance after learning several other tasks. Models that were built using classical sequential updating techniques, in conjunction with the addition of new tasks, clearly demonstrated forgetting of previously learned tasks. The addition of new data distributions led to a strong decrease in accuracy on previously learned tasks, especially in steady environments with strong distributional changes.

The rehearsal mechanism allows the model to retain former memory across tasks and along training processes. In a high-degree task similarity situation, the forgetting rate is practically zero, evidencing the model's knowledge consolidation capacity. Additionally, the data drift's impact on forgetting rates indicated the existence of a 'sweet spot' of slighter distribution changes. In a lower range, the model was observed to learn better, while at a higher range the model learned worse. That said, the tested framework was able to stride through the challenges while maintaining better retention rates than the models with no rehearsal mechanism.

The framework was also tested against other continual learning strategies to assess its ability to combat catastrophic forgetting. Figure 6 shows the forgetting rates across various continual learning strategies.

Figure 6. Forgetting Rate Comparison Across Continual Learning Methods

The proposed design framework was better than all the other models tested, including EWC and replay-based systems.

3.3 Computational Efficiency and Memory Overhead

Calculating the system's ability to learn continuously is important when deploying it in real life situations. Most case studies about retraining techniques involve datasets that are infinite in scope. They include retraining periods that are too lengthy and result in high computational and storage costs. Our proposed framework solves this problem because it only requires the storage of a small memory buffer that contains representative samples of the previously learned tasks. Our experiment shows that memory overhead does not change during the cumulative process of memory learning. This is because, within the buffer, the size of the data is limited and it is updated in a dynamic manner.

The Incremental Update mechanism is the most efficient in terms of training. This mechanism allows the model to learn from a newly introduced data stream with minimal effort and without needing to fully retrain the model. This approach is considerable when compared to methods that involve extensive retraining. The rehearsal mechanism does introduce a small amount of computational overhead due to the repeated samples, but in the grand scheme of things, this cost is very small when considering the advantages of enhanced knowledge retention, and the prevention of memory loss. The data in the proposed framework shows that it is possible to retain predictive stability while remaining computationally efficient, and memory usage remains low. This data convinces us that the framework can be used in the most dynamic systems without hindrances. The data in Table 5 shows our proposed framework in comparison to other learning approaches in terms of training time, memory usage, and scalability.

Table 5. Computational Efficiency Comparison

Model	Training Time per Task (s)	Memory Usage (MB)	Scalability	Remarks
Sequential Training (Baseline)	85	520	Low	Requires full retraining with accumulated data
Elastic Weight Consolidation	62	410	Medium	Regularization constraints increase training complexity
Experience Replay	55	460	Medium	Memory replay improves retention but increases storage
Proposed Continual Learning Framework	42	290	High	Incremental updates with compact rehearsal memory

Statistical significance testing (paired t-test) confirms that the performance improvements of the proposed framework over baseline methods are significant ($p < 0.01$).

3.4 Discussion

The incremental update mechanism proposed in the continual learning framework features superior computational efficiency in stream environments, compared with the other traditional frameworks. In the continual learning framework proposed, with the use of incremental update mechanisms, training time for each task is minimized, and the use of compact rehearsal memory minimizes the size and storage memory to allow the framework to operate without breaching computational and memory constraints adaptable to long-term streaming environments.

On the other hand, this research holds that continual learning, within the defined constraints, is the equilibrium point of knowledge stasis and adaptive plasticity. We contrast this with the traditional frameworks wherein memory retention is viewed as an isolated, secondary feature, whereas, in our case, retention-driven adaptation is posited to be the primary mechanism for fortifying the long-term equilibrium of learning. This research contributes to the early stages of advanced intelligent systems where continual learning is a requisite, alongside the ability to adjust to new data distributions without losing reliability.

3.4.1 Effectiveness of Memory-Based Rehearsal

The memory rehearsal mechanism mitigating the catastrophic forgetting phenomenon is one of the highlighted features in the proposed framework of the study. Active and reinforced retention of the knowledge of the past is achieved via the mechanism of storing and replaying the selected samples of the past essential tasks, and it prevents the model of focusing on a distribution of the most current tasks alone.

The strategies also show flexibility in dealing with different degrees of task similarity. If tasks are similar in terms of certain features, the replay mechanism permits the model to invoke similar tasks and knowledge, enhancing general task and knowledge performance. On the other hand, if tasks are distributed differently, the memory buffer keeps separate representations of previously stored knowledge.

Memory rehearsal systems ease the challenges of maintaining knowledge in continual learning systems. Thus, memory rehearsal systems should be integrated into adaptive AI systems.

3.4.2 Impact of Data Drift and Task Similarity

The stability in learning is affected by task similarity and data drift. Gradual change in data leads to the model adjusting and learning to cope with the data, facilitating easy task transitions. On the contrary, frequent change in data leads to poor performance in continual learning systems.

There is evidence that the proposed system works in both instances. However, performance is significantly enhanced when the data is changed gradually in the proposed system. This illustrates the need for understanding data during the development of continual learning systems.

Fraud detection, intrusion detection, recommendation systems, and other examples have real world variability associated with their data distributions. A robust continual learning framework would need to address variable data distributions in both gradual and sudden ways.

3.4.3 Practical Implications for Adaptive AI Systems

The results of this study provide several important practical implications for the development of adaptive AI systems. In the case of adaptive AI systems, traditional machine learning adaptive to the changing environment involves a costly and inefficient retraining system. A continual learning framework is a practical alternative since the model can be altered while maintaining the previous knowledge.

This framework proposal supports the notion that, in the case of incremental learning and memory rehearsal, long term stability performance can be achieved without access to the full, historical data set. This is incredibly applicable to limitless large-scale data streams with computational and storage resource constraints.

The framework's modular structure further supports integration with different machine learning models and application areas. The same applies to smart surveillance, financial intelligence, and automated reasoning systems.

3.4.4 Limitations and Future Research Directions

While this study has its positive aspects, there are still many downsides that need to be explained, the first of which is that the rehearsal mechanism depends on the contents of the memory buffer, and so if the stored sample selected does not sufficiently represent the previous tasks, the system will still suffer from partial forgetting.

Regarding the memory buffer, its size creates a limitation of knowledge retention versus storage, which, in simpler terms, Large buffers provide more historical data to be retained and need more memory, so this will be a problem in the storage dimension. In terms of research and advancing this knowledge, may include some level of sophistication, such that the memory size is allocated independently based on the nature of the tasks and the intensity of data drift.

The existing framework is mainly based on the principles of supervised learning, so the proposed framework will demand alterations to the rehearsal and update mechanisms to support the learning models. Future research may consider the combination of embedded learning and self-supervised representation learning within evolving data streams to enhance the framework proposed for the learning.

Conclusion

4.1 Conclusion

This study examines the challenges associated with deploying machine learning systems in dynamic environments where data distributions evolve over time. Traditional machine learning systems perform poorly in real world situations where they must deal with streams of data at varying intervals, and patterns, because they are designed with an unspoken assumption of a collection of data that is temporally unchanging. In these types of environments, one of the greatest dangers is catastrophic forgetting, where a model is unable to recall and apply previously acquired knowledge when it is retrained in response to new data. Overcoming catastrophic forgetting is one of the greatest challenges that adaptive AI systems must be able to achieve if they are to be able to perform the same tasks reliably over an extended period of time.

In response to these challenges, this study develops a new modular continual learning framework that combines the incremental learning and rehearsal memory techniques. The goal of the new approach is to allow a model to adjust to new data, after it has been trained using that data to a point where it recalls target knowledge from a definable range of previously acquired data. The novelty of this approach is the construction of a memory buffer of training data and model behaviors that helps to prevent catastrophic forgetting of target knowledge, while allowing the model to respond adaptively to new information.

Experimental evaluation conducted using sequential learning benchmarks demonstrates that the proposed framework maintains stable predictive performance across evolving tasks. The framework outperforms all other traditional approaches to sequential learning and activity recording without incorporating rehearsal strategies. For the study, all traditional methods were employed, and the framework was able to reduce the rate of learning loss to an all-time low while maintaining high accuracy on tasks that had already been learned. The study also reinforced the earlier knowledge retention and highlighted the need to include representative historical data during model updates.

Besides enhancements in predictive performance, the solution outlined here offers tangible benefits in computational time and memory use. Also, the outlined solution uses less memory compared to the traditional retraining strategies that use the entire dataset accessed multiple times. Thus, it can be deployed in situ for long periods in data and compute constrained environments. This use case is also applicable for the incremental mechanism that reduces the training burden, because it allows the model to adjust to new data streams without complete retraining.

This research has also sought to explore task similarity and data drift and its impact on the stability of the continual learning process. It was observed that gradual shifts in data distribution are beneficial for the integration of knowledge, while sudden disruptions are detrimental to maintenance of the stability of the learning process. However, the framework presented in the study, in comparison to other frameworks, shows stability under the mentioned shifts, and suggests that memory-based rehearsal improves the stability of the learning process.

This study provides the continual learning community with a scalable framework that helps solve the problem of balancing knowledge retention, adaptability, and storage efficiency. This study captures the essence of designing machine learning applications that operate in a continuously changing data environment, an important aspect in the majority of real world usage, like in cyber-attack detection, personalized recommendations, business intelligence, and decision support systems. The framework adds to the knowledge in the area of continual learning that combining incremental learning with a rehearsal memory of limited capacity is practical and scalable for alleviating catastrophic forgetting in an environment of adaptive learning over extended periods.

4.2 Future Work

The continual learning framework provides a solid foundation for alleviating catastrophic forgetting and supporting responsive learning in changing data streams, yet a lot of work is needed to further its development.

To improve the efficiency of the rehearsal mechanism, more advanced memory management strategies can be suggested. Currently, the framework captures representative samples and stores them in a memory buffer of fixed size. While preserving important memories of previous tasks, other more advanced sample selection methods can be used, like diversity-based sampling, memory compression via clustering, or importance-weighted selection of memories; retention of knowledge can also be increased, and storage can be reduced.

Second, the integration of self-supervised or representation learning techniques may improve the generalization capability of continual learning systems. Potentially, models could learn feature representations that galvanize them to adapt to changing environments, lessening the need for labeled data. Combining continual learning with self-supervised learning paradigms could therefore lead to more scalable adaptive intelligence systems.

Third, future studies could investigate the application of the proposed framework in large-scale real-world data streams. Although benchmark data sets offer a simplified setting for the assessment of algorithms in continual learning, real-world data sets have more complicated patterns that include non-stationary distributions, temporal irregular task boundaries, and a variety of feature spaces. Deploying the framework in real-world settings, such as network intrusion detection, industrial sensor monitoring, and analysis of financial transactions, can provide more practical insights.

Additionally, a further viable approach is the unification of continual learning frameworks with distributed or federated learning systems. In a variety of contemporary scenarios, data streams originate from multi-regional dispersed sources, while the centralized storage of such data is not feasible, or is restricted due to privacy concerns. The fusion of continual learning and federated learning frameworks may lead to the emergence of systems with distributed adaptive collaborative intelligence.

Furthermore, future research may be directed toward the theoretical aspects of the stability of continual learning and may include the formal analysis of the dynamics of forgetting, the stability of the system's optimization depending on a sequence of updates, and the relation of the similarity of tasks to the transfer of knowledge. These theoretical studies can help clarify the mechanisms of long-term learning in the adaptive frameworks of AI.

Memory’s adaptive strategies, where the task complexity and the data drift determine how the buffer limits are set, present another promising direction for research. Such strategies would help in the long-term retention of knowledge.

The importance of continual learning lies in the opportunity it presents for the development of intelligent systems able to operate in dynamic environments. Along with tackling the issues associated with the static shortcomings of ML systems, this research aims to propose additional refinements to retention and adaptation mechanisms. This, in turn, relates to the overarching aim of developing more advanced artificial intelligence systems that can operate within the boundaries of the system, regardless of the time span of the system’s functioning.

Acknowledgements

The researcher did not receive any funding for this study, and the results have not been published in any other sources. Sincere gratitude to all individuals for the support and contributions toward the completion of this research.

References

A H, N. F., Khraisat, A., S P, S. I., & Li, G. (2025). Adaptive memory replay for network intrusion detection: Tackling data drift and catastrophic forgetting. Computer Networks, 272(3), 111712. https://doi.org/10.1016/j.comnet.2025.111712

Cacciarelli, D., Kulahci, M., Gama Davide Cacciarelli, J. B., & Kulahci muku, M. (2023). Active learning for data streams: a survey. Machine Learning 2023 113:1, 113(1), 185–239. https://doi.org/10.1007/s10994-023-06454-2

Calderon, G., del Campo, G., Saavedra, E., & Santamaría, A. (2023). Monitoring Framework for the Performance Evaluation of an IoT Platform with Elasticsearch and Apache Kafka. Information Systems Frontiers 2023 26:6, 26(6), 2373–2389. https://doi.org/10.1007/s10796-023-10409-2

Chen, T., Chen, X., Chen, W., Heaton, H., Liu, J., Wang, Z., & Yin, W. (2022). Learning to Optimize: A Primer and A Benchmark. Journal of Machine Learning Research, 23(189), 1–59. http://jmlr.org/papers/v23/21-0308.html

Chergui, M., Nagano, A., & Ammoumou, A. (2024). Toward an adaptive learning system by managing pedagogical knowledge in a smart way. Multimedia Tools and Applications 2024 84:24, 84(24), 27777–27793. https://doi.org/10.1007/s11042-024-20207-w

Chhajer, P., Shah, M., & Kshirsagar, A. (2022). The applications of artificial neural networks, support vector machines, and long–short term memory for stock market prediction. Decision Analytics Journal, 2(1), 100015. https://doi.org/10.1016/j.dajour.2021.100015

Deering, K., Brimblecombe, N., Matonhodze, J. C., Nolan, F., Collins, D. A., & Renwick, L. (2023). Methodological procedures for priority setting mental health research: a systematic review summarising the methods, designs and frameworks involved with priority setting. Health Research Policy and Systems 2023 21:1, 21(1), 64-. https://doi.org/10.1186/s12961-023-01003-8

Faber, K., Corizzo, R., Sniezynski, B., & Japkowicz, N. (2024). Lifelong Continual Learning for Anomaly Detection: New Challenges, Perspectives, and Insights. IEEE Access, 12, 41364–41380. https://doi.org/10.1109/ACCESS.2024.3377690

Fang, J., Zhu, Z., Li, S., Su, H., Yu, Y., Zhou, J., & You, Y. (2023). Parallel Training of Pre-Trained Models via Chunk-Based Dynamic Memory Management. IEEE Transactions on Parallel and Distributed Systems, 34(1), 304–315. https://doi.org/10.1109/TPDS.2022.3219819

Fragkoulis, M., Carbone, P., Kalavri, V., & Katsifodimos, A. (2023). A survey on the evolution of stream processing systems. The VLDB Journal 2023 33:2, 33(2), 507–541. https://doi.org/10.1007/s00778-023-00819-8

Han, Y. nan, & Liu, J. wei. (2024). Adaptive instance similarity embedding for online continual learning. Pattern Recognition, 149, 110238. https://doi.org/10.1016/j.patcog.2023.110238

Hou, J., Cosma, G., & Finke, A. (2025). Advancing continual lifelong learning in neural information retrieval: Definition, dataset, framework, and empirical evaluation. Information Sciences, 687, 121368. https://doi.org/10.1016/j.ins.2024.121368

Iman, N. (2025). Différance and Data Drift: The Trace of Change in Machine Learning. Philosophy & Technology 2025 38:3, 38(3), 98-. https://doi.org/10.1007/s13347-025-00936-y

Jung, E., Lim, R., & Kim, D. (2022). A Schema-Based Instructional Design Model for Self-Paced Learning Environments. Education Sciences 2022, Vol. 12, 12(4). https://doi.org/10.3390/educsci12040271

Khosravi, H., Olajire, T., Raihan, A. S., & Ahmed, I. (2024). A data driven sequential learning framework to accelerate and optimize multi-objective manufacturing decisions. Journal of Intelligent Manufacturing 2024 35:8, 35(8), 4087–4112. https://doi.org/10.1007/s10845-024-02337-y

Liang, X., Tang, J., Zhong, Y., Gao, B., Qian, H., & Wu, H. (2024). Physical reservoir computing with emerging electronics. Nature Electronics 2024 7:3, 7(3), 193–206. https://doi.org/10.1038/s41928-024-01133-z

Santoro, D., Ciano, T., & Ferrara, M. (2024). A comparison between machine and deep learning models on high stationarity data. Scientific Reports 2024 14:1, 14(1), 19409-. https://doi.org/10.1038/s41598-024-70341-6

Shaheen, K., Hanif, M. A., Hasan, O., & Shafique, M. (2022). Continual Learning for Real-World Autonomous Systems: Algorithms, Challenges and Frameworks. Journal of Intelligent & Robotic Systems 2022 105:1, 105(1), 9-. https://doi.org/10.1007/s10846-022-01603-6

Shen, L., Sun, Y., Yu, Z., Ding, L., Tian, X., & Tao, D. (2024). On Efficient Training of Large-Scale Deep Learning Models. ACM Computing Surveys, 57(3). https://doi.org/10.1145/3700439

Shi, H., Xu, Z., Wang, H., Qin, W., Wang, W., Wang, Y., Wang, Z., Ebrahimi, S., & Wang, H. (2025). Continual Learning of Large Language Models: A Comprehensive Survey. ACM Computing Surveys, 58(5). https://doi.org/10.1145/3735633

Talaei Khoei, T., Ould Slimane, H., & Kaabouch, N. (2023). Deep learning: systematic review, models, challenges, and research directions. Neural Computing and Applications 2023 35:31, 35(31), 23103–23124. https://doi.org/10.1007/s00521-023-08957-4

Wang, L., Zhang, X., Su, H., & Zhu, J. (2024). A Comprehensive Survey of Continual Learning: Theory, Method and Application. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(8), 5362–5383. https://doi.org/10.1109/TPAMI.2024.3367329

Wang, Z., Yang, E., Shen, L., & Huang, H. (2025). A Comprehensive Survey of Forgetting in Deep Learning Beyond Continual Learning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 47(3), 1464–1483. https://doi.org/10.1109/TPAMI.2024.3498346

Wickramasinghe, B., Saha, G., & Roy, K. (2024). Continual Learning: A Review of Techniques, Challenges, and Future Directions. IEEE Transactions on Artificial Intelligence, 5(6), 2526–2546. https://doi.org/10.1109/TAI.2023.3339091

Wu, B., Ding, Z., & Huang, J. (2026). A Review of Continual Learning in Edge AI. IEEE Transactions on Network Science and Engineering. https://doi.org/10.1109/TNSE.2026.3657652

Yang, X., Yu, H., Gao, X., Wang, H., Zhang, J., & Li, T. (2024). Federated Continual Learning via Knowledge Fusion: A Survey. IEEE Transactions on Knowledge and Data Engineering, 36(8), 3832–3850. https://doi.org/10.1109/TKDE.2024.3363240