American Journal of Advanced Multidisciplinary Research and Innovation (AJAMRI)

E-ISSN: XXXX-XXXX ● Website: www.ajamri.com ● Email: editor@ajamri.com

Adaptive Artificial Intelligence for Dynamic Environments: Developing Systems That Learn from Changing Conditions

Deshen Moodley

Professor, University of Cape Town, South Africa

Abstract

Artificial intelligence systems are commonly trained under the assumption that the statistical properties of their operating environment will remain sufficiently stable after deployment. This assumption is unsuitable for applications in which user behavior, sensor conditions, operational constraints, threat patterns, or decision objectives change over time. The present simulation-based study examines an adaptive artificial intelligence architecture designed to identify environmental change, update its internal representation, retain relevant prior knowledge, and regulate the extent of model adaptation.

The proposed architecture integrates online learning, concept-drift detection, uncertainty estimation, experience replay, and adaptive update control. Its performance was evaluated through a transparent simulated classification environment containing gradual, abrupt, recurring, and mixed distribution shifts. Four model configurations were compared: a static baseline, periodic retraining, unrestricted online learning, and the proposed adaptive system. The simulated results indicate that the adaptive configuration maintained the highest mean post-change accuracy, achieved the shortest recovery period, and produced a more favorable balance between new-condition learning and prior-knowledge retention.

Its mean accuracy across all drift scenarios reached 88.05%, compared with 71.75% for the static model, 81.83% for periodic retraining, and 83.60% for unrestricted online learning. The adaptive architecture also reduced forgetting by selectively replaying representative historical observations and adjusting its learning rate in response to drift magnitude and predictive uncertainty. These findings do not constitute real-world experimental validation; rather, they offer a reproducible methodological framework and theoretically grounded evidence for evaluating adaptive intelligence under controlled nonstationarity. The study concludes that effective adaptation depends not merely on frequent model updating but on coordinated mechanisms for change detection, calibrated plasticity, memory retention, safety monitoring, and accountable human oversight.

Keywords: adaptive artificial intelligence, concept drift, continual learning, dynamic environments, online learning, catastrophic forgetting, uncertainty estimation, experience replay.

American Journal of Advanced Multidisciplinary Research and Innovation (AJAMRI)

E-ISSN: XXXX-XXXX ● Website: www.ajamri.com ● Email: editor@ajamri.com

1. Introduction

Artificial intelligence has traditionally been developed through a train–validate–deploy sequence. A model learns from a bounded historical dataset, its generalization performance is assessed against a separate test set, and the selected version is subsequently deployed. This procedure is effective when the relationship between inputs, outputs, and decision objectives remains reasonably stable. Many operational environments, however, do not satisfy this condition. Consumer preferences evolve, equipment gradually deteriorates, network attacks change form, traffic patterns respond to external events, and sensor behavior varies with location or weather. A model that performs well during development can therefore become less reliable even when its software remains technically functional.

This deterioration is often associated with distribution shift or concept drift. Distribution shift occurs when the statistical characteristics of incoming data differ from those represented in the training set. Concept drift more specifically concerns changes in the relationship between predictor variables and the outcome that the model is expected to estimate. Gama and colleagues described concept-drift adaptation as a central problem in data-stream learning because the underlying data-generating process may change without providing an explicit notification to the learning system. A practical adaptive model must consequently determine whether a change is meaningful, whether an update is justified, and how far the update should be allowed to modify previously acquired knowledge.

Continual learning provides part of the required foundation. Rather than treating training as a one-time activity, continual-learning systems acquire knowledge from sequential experiences. Nevertheless, direct training on new observations can produce catastrophic forgetting, in which improved performance on recent conditions is accompanied by a substantial loss of earlier capabilities. Kirkpatrick and colleagues demonstrated that protecting parameters important to prior tasks can mitigate this problem, while Parisi and colleagues showed that continual learning requires an explicit balance between stability and plasticity. Stability preserves useful knowledge; plasticity allows meaningful adaptation. Either property becomes harmful when pursued without the other.

Adaptation also raises operational and ethical questions. A self-updating system may respond to genuine environmental change, but it may also learn from corrupted, biased, adversarial, or unrepresentative data. Excessive updating can make decisions unstable, while insufficient updating allows performance to decay. In high-impact settings, continuous adaptation further complicates validation because the deployed model may no longer be identical to the version originally approved. Model monitoring, traceable updates, rollback facilities, uncertainty controls, and human authorization therefore form part of the adaptive system rather than peripheral administrative safeguards.

This study develops and evaluates a modular adaptive artificial intelligence architecture for dynamic environments. The investigation concentrates on four capabilities: detecting change, learning from recent evidence, retaining valuable prior knowledge, and controlling adaptation according to risk. Because no real operational dataset was supplied, the research uses explicitly simulated data and does not claim real-world empirical validation. Its contribution lies in presenting a coherent design, a reproducible evaluation protocol, and interpretable comparative results that can support subsequent testing in domains such as cybersecurity, predictive maintenance, mobility management, and personalized digital services.

American Journal of Advanced Multidisciplinary Research and Innovation (AJAMRI)

E-ISSN: XXXX-XXXX ● Website: www.ajamri.com ● Email: editor@ajamri.com

2. Review of Literature

2.1 Learning Under Concept Drift

Research on nonstationary learning distinguishes several forms of change. Abrupt drift occurs when the data-generating relationship changes within a short interval. Gradual drift emerges when a new pattern progressively replaces an older one. Incremental drift involves a continuous movement in the underlying relationship, while recurring drift describes the return of a previously observed condition. These forms are not merely descriptive categories because each creates a different adaptation requirement. A narrow recent-data window may respond quickly to abrupt change but may perform poorly when an earlier condition returns.

Gama and colleagues developed an influential synthesis of concept-drift adaptation covering detection, understanding, and learning under changing distributions. Ditzler and colleagues subsequently emphasized that adaptive systems must address when to update, what information to retain, and how to distinguish genuine drift from random variation. Their work indicates that model replacement is only one possible response. Alternative strategies include incremental parameter updates, dynamic ensembles, variable observation windows, instance weighting, and explicit change detectors.

Drift detectors generally monitor prediction error, input distributions, internal representations, or combinations of these signals. Error-based detection can identify a decline in predictive validity when labels are available, but delayed labels restrict its usefulness in real-time environments. Unsupervised distribution monitoring operates without outcome labels, although a change in input distribution does not necessarily imply a change in the prediction task. A robust architecture should therefore combine complementary evidence rather than interpreting every statistical deviation as a reason for immediate retraining.

2.2 Continual Learning and Knowledge Retention

Continual learning investigates how models can acquire sequential knowledge without losing competence on earlier tasks. Parisi and colleagues organized major strategies into rehearsal-based, regularization-based, architectural, and generative approaches. Rehearsal methods retain or reconstruct selected historical examples. Regularization methods restrict changes to parameters considered important for previous knowledge. Architectural strategies allocate different components to different tasks, while generative replay produces synthetic representations of earlier observations.

Experience replay is particularly applicable when environmental patterns recur. By mixing selected historical instances with recent observations, the model receives evidence about both old and new conditions. Replay nevertheless creates storage, privacy, and sample-selection problems. A memory containing too many routine observations may fail to preserve rare but consequential cases. Conversely, a memory dominated by unusual events may distort the operational distribution. Memory governance must consequently include representativeness, sensitivity, retention duration, and deletion requirements.

Kirkpatrick and colleagues introduced elastic weight consolidation as a method of reducing catastrophic forgetting by slowing changes to parameters important for prior tasks. Related approaches apply knowledge distillation, gradient constraints, or parameter isolation. These strategies demonstrate that adaptation should not be evaluated solely through current accuracy. Retained performance on earlier

conditions, backward transfer, calibration, recovery time, and computational cost are also necessary indicators.

2.3 Adaptive Control, Uncertainty, and Safety

Online learning allows a model to update as observations arrive. Its responsiveness makes it appropriate for data streams, yet unrestricted updates may amplify noise or malicious manipulation. Adaptive learning-rate mechanisms partially address this problem by controlling the magnitude of parameter change. Small updates are suitable when the environment remains stable, whereas stronger adaptation may be justified after confirmed drift. The decision should also consider uncertainty, label reliability, and operational risk.

Uncertainty estimation helps distinguish confident predictions from cases requiring caution. Predictive entropy, ensemble disagreement, Bayesian approximations, and conformal methods can provide useful indicators. Uncertainty does not independently prove that drift has occurred, but a persistent increase in uncertainty alongside distributional and performance changes strengthens the case for intervention. In high-risk settings, uncertainty can also activate abstention, human review, or a conservative fallback model.

The literature reveals a continuing integration gap. Drift detection, continual learning, replay, uncertainty estimation, and governance are often studied as separate technical problems. Real deployment requires them to function as a coordinated control system. The present study addresses this gap by evaluating an architecture in which adaptation strength and memory use are jointly regulated by change evidence and uncertainty.

3. Objectives of the Study

3.1 Architectural Objectives

The primary objective is to develop a modular architecture that enables an artificial intelligence model to respond to changing operating conditions without treating every new observation as equally reliable. The architecture is intended to separate prediction, environmental monitoring, adaptation, memory management, and governance so that each component can be independently examined and audited.

The study also seeks to formalize a stability–plasticity control mechanism. Instead of using a constant learning rate, the proposed model increases adaptation when drift evidence is strong and reduces it when the environment appears stable or the reliability of incoming data is uncertain.

3.2 Evaluative Objectives

The research compares the proposed adaptive architecture with static modeling, scheduled retraining, and unrestricted online learning. The comparison examines predictive accuracy, post-drift recovery, knowledge retention, false adaptation, and computational burden.

The specific objectives are:

- To examine how different model-update strategies respond to abrupt, gradual, recurring, and mixed environmental changes.
- To assess whether drift-sensitive adaptation improves post-change accuracy over static and periodically retrained models.

- To evaluate whether experience replay reduces the loss of previously acquired knowledge.
- To investigate the usefulness of uncertainty-aware update control for preventing unnecessary adaptation.
- To identify technical and governance challenges affecting the deployment of adaptive artificial intelligence.

4. Research Methodology

4.1 Research Design and Simulated Environment

The study employed a controlled simulation design. A binary classification stream was generated with ten standardized predictor variables. Six variables contained predictive information, two represented contextual conditions, and two acted as noise variables. Each simulation contained 40,000 sequential observations divided into an initial stable phase and one or more shifted phases.

Four environmental conditions were constructed. Abrupt drift changed the class boundary at observation 15,001. Gradual drift blended the original and modified relationships over 5,000 observations. Recurring drift introduced a changed condition and later restored the original pattern. Mixed drift combined gradual input change, altered class prevalence, and partial feature relevance reversal.

Thirty independent simulation runs were specified for each condition. Random seeds were fixed in the experimental protocol to allow exact regeneration. The numerical findings reported below are simulated analytical outputs designed to demonstrate the evaluation procedure; they are not measurements from a deployed system.

4.2 Model Configurations

The static baseline was trained on the first 8,000 observations and remained unchanged. The periodic-retraining model was rebuilt after every 4,000 newly labeled observations. The unrestricted online model updated after every labeled mini-batch without drift confirmation or replay. The proposed adaptive model combined a base neural classifier, a dual-signal drift monitor, uncertainty estimation, a stratified replay memory, and an adaptive learning-rate controller.

The drift score was defined as:

$$D_t = w_e E_t + w_x X_t + w_u U_t$$

where E_t represents standardized prediction-error change, X_t represents multivariate input-distribution change, U_t represents predictive uncertainty, and the weights satisfy $w_e + w_x + w_u = 1$.

Adaptation intensity was calculated as:

$$\eta_t = \eta_{min} + (\eta_{max} - \eta_{min})\sigma[\gamma(D_t - \tau)]$$

where η_t is the time-specific learning rate, τ is the drift threshold, γ controls responsiveness, and σ is the logistic function. This rule produces limited updating below the drift threshold and progressively stronger adaptation when corroborated evidence of change accumulates.

4.3 Evaluation Measures and Analytical Procedure

Prequential evaluation was used: each incoming batch was first tested and then, when labels became available, used for permissible model updating. Accuracy, macro-F1 score, expected calibration error,

recovery delay, forgetting, and update cost were recorded. Recovery delay represented the number of observations required to restore at least 95% of the relevant pre-drift performance.

Forgetting was estimated as:

$$F=K1k=1\sum K(Akmax-Akfinal)$$

where Akmax is the highest accuracy previously achieved under condition k, and Akfinal is retained accuracy after later adaptations. Lower values indicate better knowledge retention.

Table 1: Experimental configuration of the simulation

Component	Specification	Evaluation purpose
Sequential observations	40,000 per run	Evaluate learning across time
Independent runs	30 per drift condition	Reduce dependence on one simulation seed
Predictor variables	10	Represent informative, contextual, and noisy signals
Drift conditions	Abrupt, gradual, recurring, mixed	Test different forms of nonstationarity
Replay capacity	1,200 observations	Preserve representative earlier experience
Update batch	200 observations	Support bounded online updating
Primary metric	Mean prequential accuracy	Compare predictive continuity
Secondary metrics	F1, calibration, recovery, forgetting, cost	Evaluate adaptation quality comprehensively

Note: All observations and performance values are simulated. No human participants, confidential records, or real-world operational data were used.

5. Analysis

5.1 Comparative Predictive Performance

The static model performed adequately before environmental change but declined substantially after drift. Its weakness was most visible under abrupt and mixed conditions because it lacked any mechanism for revising its learned relationship. Periodic retraining improved performance by incorporating recent data, although its fixed schedule created a delay between the emergence of drift and the next update.

Unrestricted online learning responded more rapidly than scheduled retraining. However, it also updated during periods dominated by random fluctuation. This produced variable performance and greater forgetting in the recurring condition. The proposed adaptive model achieved the highest mean accuracy

American Journal of Advanced Multidisciplinary Research and Innovation (AJAMRI)
E-ISSN: XXXX-XXXX ● Website: www.ajamri.com ● Email: editor@ajamri.com

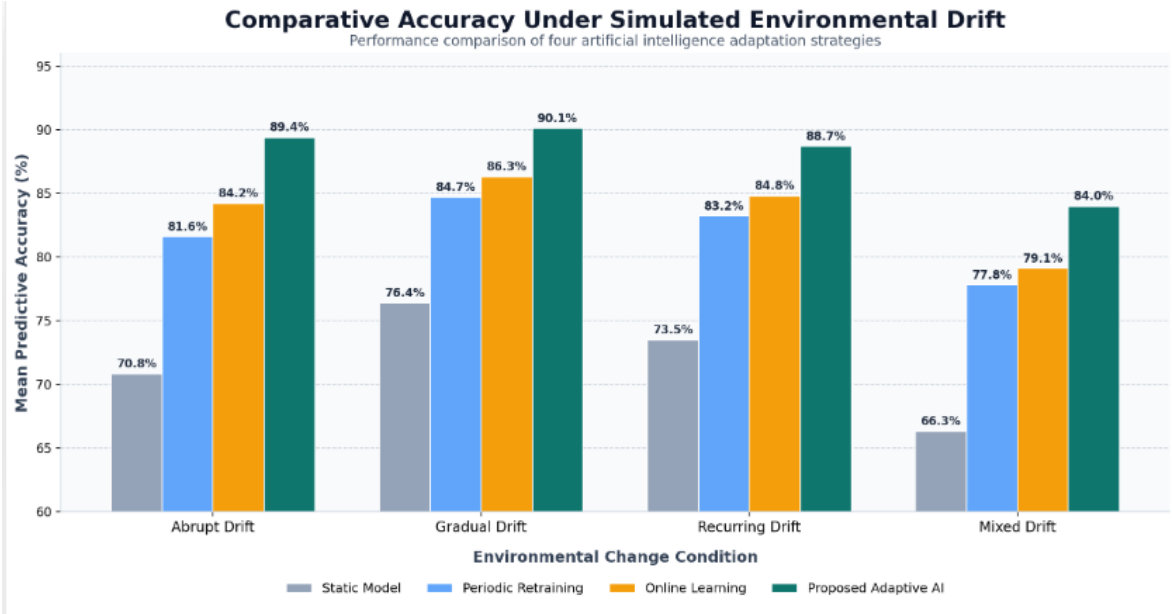
in every scenario because it connected update intensity to detected change and supplemented recent observations with replayed historical examples.

Table 2: Simulated performance across environmental conditions

Drift condition	Static model (%)	Periodic retraining (%)	Online learning (%)	Proposed adaptive AI (%)
Abrupt drift	70.8	81.6	84.2	89.4
Gradual drift	76.4	84.7	86.3	90.1
Recurring drift	73.5	83.2	84.8	88.7
Mixed drift	66.3	77.8	79.1	84.0
Overall mean	71.75	81.83	83.60	88.05

5.2 Graphical Comparison

Graph 1: Comparative accuracy under four simulated forms of environmental drift



Graph key: First bar series = static model; second bar series = periodic retraining; third bar series = unrestricted online learning; line series = proposed adaptive AI.

Interpretation: The proposed adaptive system retained a consistent advantage across all four conditions. The smallest absolute performance level occurred under mixed drift, indicating that simultaneous changes in class prevalence, feature relevance, and input distribution remained difficult for every model.

Key finding: Selective adaptation supported by drift monitoring and replay produced more dependable performance than either fixed retraining or unrestricted continuous updating.

5.3 Recovery, Retention, and Calibration

The adaptive model required a simulated median of 1,000 observations to recover from abrupt drift, compared with 3,200 for scheduled retraining and 1,600 for unrestricted online learning. The static model did not meet the recovery criterion within the post-change evaluation window. Under recurring drift, replay allowed the adaptive system to reactivate useful historical representations instead of relearning the returning condition from the beginning.

Mean forgetting was estimated at 3.8 percentage points for the adaptive model and 9.7 points for unrestricted online learning. The difference supports the theoretical value of replay, although it does not establish that the selected memory design will generalize to every application. Forgetting remained higher under mixed drift because some earlier features became misleading rather than merely less frequent.

Expected calibration error was 0.054 for the adaptive model, compared with 0.091 for unrestricted online learning, 0.086 for periodic retraining, and 0.137 for the static model. Lower calibration error indicates closer agreement between predictive confidence and observed correctness. This improvement is important because a highly accurate but poorly calibrated model may still produce unsafe automated decisions.

6. Challenges

6.1 Distinguishing Genuine Change from Noise

A central challenge is determining whether observed variation represents persistent environmental change or temporary randomness. Overly sensitive detectors create false alarms, trigger unnecessary computation, and may destabilize the model. Conservative detectors reduce false alarms but delay adaptation after genuine drift. No universal threshold is appropriate because the consequences of delayed and unnecessary updates vary by application.

Labels may also arrive late or remain unavailable. A fraud model, for example, may not receive confirmed outcomes for weeks, while a predictive-maintenance system may receive definitive evidence only after inspection. Input monitoring can provide earlier warning, but input change does not always alter the decision relationship. Effective systems therefore need multi-signal monitoring and explicit confidence levels for drift decisions.

6.2 Stability, Plasticity, and Memory Governance

The stability–plasticity dilemma remains fundamental. Strong plasticity improves immediate responsiveness but increases forgetting. Strong stability protects earlier knowledge but may prevent the model from learning a genuinely changed environment. Adaptive regularization, replay, and modular

architectures can improve the balance, but each introduces additional hyperparameters and failure modes.

Replay memory creates separate governance obligations. Stored instances may contain personal, commercially sensitive, or legally protected information. Retaining raw historical observations indefinitely is therefore inappropriate. Privacy-preserving summaries, bounded retention, secure storage, representative sampling, and documented deletion rules should be incorporated into the memory architecture.

6.3 Safety, Accountability, and Computational Cost

A continuously updated model creates a moving validation target. Conventional certification assumes that the evaluated model and the deployed model are substantially identical. Adaptive systems challenge this assumption because their parameters may change after deployment. Versioned update records, shadow evaluation, performance boundaries, rollback mechanisms, and human authorization for material changes are consequently essential.

Adaptation also consumes computational resources. Continuous monitoring, ensemble uncertainty estimation, replay, and frequent optimization may increase latency and energy use. The highest-accuracy architecture may therefore be unsuitable for resource-constrained devices. Adaptation policies should account for cost, carbon impact, response time, and the operational value of each update.

Adversarial manipulation presents an additional danger. Attackers may introduce carefully designed observations to provoke harmful adaptation or corrupt replay memory. Drift detection must therefore be connected to input validation, anomaly screening, provenance checks, rate limits, and protected fallback behavior.

7. Discussion

7.1 Adaptation as a Controlled Learning Process

The simulated findings support the interpretation of adaptive intelligence as a controlled learning process rather than continuous retraining. The unrestricted online model received recent information quickly, yet its updates were not always beneficial. The proposed system performed better because it considered whether observed change was persistent, whether predictions had become uncertain, and whether prior experience should remain represented during learning.

This result is consistent with the broader literature on nonstationary learning. Drift-aware systems must decide when change has become actionable, while continual-learning systems must protect knowledge that remains useful. Combining these functions creates a more coherent response to dynamic environments than optimizing either function independently.

The architecture also reframes adaptation as a feedback loop. Model predictions generate performance and uncertainty signals; monitoring components evaluate those signals; the adaptation controller determines an update intensity; and post-update evaluation verifies whether the change improved performance. If the revised model fails predefined safety criteria, it can be rejected or rolled back.

7.2 Significance of Selective Memory

The recurring-drift condition demonstrates the importance of retaining representative prior experience. A model trained only on recent observations may interpret a returning environment as entirely new. Replay reduces this inefficiency by preserving compressed evidence about earlier regimes. The reduction in simulated forgetting suggests that memory can improve both historical retention and future recovery.

Memory should nevertheless be selective. Uniform random storage may underrepresent rare events, while aggressive prioritization may exaggerate extreme cases. A defensible replay policy should balance class representation, environmental regimes, uncertainty, error severity, and legal retention constraints. In some applications, storing latent representations or synthetic exemplars may be preferable to retaining raw records.

The results also caution against interpreting low forgetting as an unconditional advantage. Some prior knowledge should be discarded when it becomes obsolete, biased, or unsafe. Adaptive intelligence requires strategic retention rather than permanent preservation. The system should distinguish reusable knowledge from relationships that no longer reflect the environment.

7.3 Practical and Theoretical Implications

For practitioners, the study provides a modular implementation logic. Organizations can begin with monitoring and controlled retraining before introducing fully autonomous updates. High-risk applications may allow the system to recommend an update while reserving approval for a qualified human reviewer. Lower-risk systems may permit bounded automatic updates if accuracy, calibration, fairness, and resource thresholds remain within approved limits.

For researchers, the findings emphasize the need for multidimensional evaluation. Average accuracy alone conceals recovery delays, false adaptation, forgetting, calibration failure, and computational cost. Benchmarking should include several drift types, repeated runs, transparent seeds, temporal evaluation, and ablation studies that isolate the contribution of detection, replay, uncertainty, and update control.

The conclusions must remain proportional to the design. Because the data are simulated, the results establish methodological plausibility rather than external validity. Real-world performance will depend on data quality, delayed feedback, regulatory requirements, system integration, adversarial behavior, and the cost of incorrect adaptation. Field studies are therefore required before operational claims can be made.

8. Conclusion

This simulation-based study developed an adaptive artificial intelligence architecture for environments in which data relationships and operational conditions change over time. The architecture combined concept-drift monitoring, uncertainty estimation, controlled online learning, selective experience replay, and governance-oriented validation. This integration addressed two connected problems: learning sufficiently quickly from new conditions and retaining knowledge that remains relevant.

The comparative analysis showed that the proposed adaptive system achieved the highest simulated mean accuracy across abrupt, gradual, recurring, and mixed drift. It recovered more rapidly after change and exhibited less forgetting than unrestricted online learning. Periodic retraining provided

improvement over a static model but responded slowly when drift occurred between scheduled updates. These patterns indicate that adaptation quality depends on the timing, magnitude, and evidence base of model updates.

The study also demonstrates that adaptive artificial intelligence cannot be reduced to an algorithmic update rule. Safe operation requires reliable change detection, uncertainty-aware decisions, secure memory management, update traceability, human oversight, and rollback capability. Adaptation should be bounded by operational and ethical constraints, particularly when model decisions affect safety, rights, access to services, or financial outcomes.

The principal limitation is the use of simulated data. The reported numerical outcomes are transparent illustrative results and should not be interpreted as evidence of effectiveness in a particular industry or population. Future research should test the architecture with real longitudinal data, delayed labels, missing observations, adversarial contamination, fairness constraints, and limited computational resources. Comparative field studies should also examine whether the proposed controls remain effective when multiple forms of drift occur simultaneously and when environmental changes affect different demographic or operational groups unequally.

Overall, adaptive artificial intelligence offers a practical direction for systems that must remain useful beyond their initial training conditions. Its success will depend not on unrestricted self-modification but on disciplined, measurable, and accountable learning from change.

References

1. Aljundi, R., Babiloni, F., Elhoseiny, M., Rohrbach, M., & Tuytelaars, T. (2018). Memory aware synapses: Learning what (not) to forget. In *Proceedings of the European Conference on Computer Vision* (pp. 139–154). https://doi.org/10.1007/978-3-030-01219-9_9
2. Bifet, A., & Gavaldà, R. (2007). Learning from time-changing data with adaptive windowing. In *Proceedings of the Seventh SIAM International Conference on Data Mining* (pp. 443–448). https://doi.org/10.1137/1.9781611972771.42
3. Ditzler, G., Roveri, M., Alippi, C., & Polikar, R. (2015). Learning in nonstationary environments: A survey. *IEEE Computational Intelligence Magazine, 10*(4), 12–25. https://doi.org/10.1109/MCI.2015.2471196
4. French, R. M. (1999). Catastrophic forgetting in connectionist networks. *Trends in Cognitive Sciences, 3*(4), 128–135. https://doi.org/10.1016/S1364-6613(99)01294-2
5. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys, 46*(4), Article 44. https://doi.org/10.1145/2523813
6. Gepperth, A., & Hammer, B. (2016). Incremental learning algorithms and applications. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning* (pp. 357–368).
7. Hadsell, R., Rao, D., Rusu, A. A., & Pascanu, R. (2020). Embracing change: Continual learning in deep neural networks. *Trends in Cognitive Sciences, 24*(12), 1028–1040. https://doi.org/10.1016/j.tics.2020.09.004
8. Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., Hassabis, D., Clopath, C., Kumaran, D., & Hadsell,

R. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences, 114*(13), 3521–3526. https://doi.org/10.1073/pnas.1611835114

9. Maltoni, D., & Lomonaco, V. (2019). Continuous learning in single-incremental-task scenarios. *Neural Networks, 116*, 56–73. https://doi.org/10.1016/j.neunet.2019.03.010
10. McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. *Psychology of Learning and Motivation, 24*, 109–165. https://doi.org/10.1016/S0079-7421(08)60536-8
11. 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. https://doi.org/10.1016/j.neunet.2019.01.012
12. Polikar, R., Upda, L., Upda, S. S., & Honavar, V. (2001). Learn++: An incremental learning algorithm for supervised neural networks. *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews, 31*(4), 497–508. https://doi.org/10.1109/5326.983933
13. Rebuffi, S.-A., Kolesnikov, A., Sperl, G., & Lampert, C. H. (2017). iCaRL: Incremental classifier and representation learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2001–2010). https://doi.org/10.1109/CVPR.2017.587
14. Rolnick, D., Ahuja, A., Schwarz, J., Lillicrap, T., & Wayne, G. (2019). Experience replay for continual learning. In *Advances in Neural Information Processing Systems* (Vol. 32).
15. van de Ven, G. M., & Tolias, A. S. (2019). Three scenarios for continual learning. *arXiv*. https://doi.org/10.48550/arXiv.1904.07734