

Article

Not peer-reviewed version

---

# Synergies between Class Incremental Learning and Machine Unlearning

---

[Chang, Qi](#)\*

Posted Date: 25 September 2024

doi: 10.20944/preprints202409.1992.v1

Keywords: Incremental Learning; Machine Unlearning; Optimization; Graph Neural Network



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

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

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

Article

# Synergies between Class Incremental Learning and Machine Unlearning

Chang Qi

**Abstract:** The convergence of Class Incremental Learning (CIL) and Machine Unlearning (MU) is a rapidly developing field in machine learning, especially relevant in adaptive and privacy-sensitive environments like finance. CIL enables models to learn new data classes over time without losing previously acquired knowledge, while MU focuses on selectively forgetting specific data to comply with privacy laws or mitigate security risks. In this paper, we examine the theoretical foundations and practical applications of both approaches, particularly in the financial domain. We explore how these two paradigms interact and complement each other, discuss key algorithms, and present examples to illustrate their applications in areas such as portfolio management, fraud detection, and data privacy. Finally, we explore challenges and potential future directions in achieving optimal synergy between CIL and MU.

**Keywords:** incremental learning; machine unlearning; optimization; graph neural network

---

## Introduction

Machine learning models are increasingly employed in various domains, from image recognition to natural language processing and autonomous systems. However, these models are typically static, requiring retraining on entire datasets whenever new data is introduced or when specific data needs to be removed. This limitation is problematic in dynamic environments where data evolves continuously, and there is a growing demand for data privacy and security. Class Incremental Learning (CIL) and Machine Unlearning address these issues by enabling models to incorporate new classes of data and forget specific data points or classes without full retraining.

CIL allows machine learning models to sequentially learn new classes while retaining knowledge of previously learned classes, thus preventing catastrophic forgetting. This is crucial in applications such as autonomous driving, where new objects or conditions need to be recognized without losing the ability to detect previously known ones. In contrast, Machine Unlearning provides the capability to remove the influence of specific data points from a trained model, a requirement increasingly driven by privacy laws such as the General Data Protection Regulation (GDPR). The ability to unlearn is essential for maintaining data privacy and ensuring that models do not retain any unwanted or erroneous information.

This paper aims to explore the potential synergies between CIL and Machine Unlearning. While these two fields have traditionally been studied in isolation, combining them could lead to more robust and flexible machine learning systems capable of both learning and forgetting as required. We will delve into the methodologies and techniques used in both fields, discuss how they can complement each other, and provide examples to illustrate their practical applications.

## Literature Review

Incremental learning as described in [1–3] aims to efficiently update novel classes with prior classes in a comprehensive model. Class Incremental Learning (CIL) is a subfield of incremental learning designed to enable models to progressively handle an increasing number of classes over time. Unlike traditional learning systems, which typically require a fixed dataset with all classes present during training, CIL allows a model to update its knowledge base when new classes are

introduced, without having to retrain the entire model from scratch. In financial applications, this is particularly important as new financial products, regulatory changes, and emerging economic factors continuously reshape the landscape.

CIL addresses two key issues inherent in its framework: catastrophic forgetting and intransigence. Catastrophic forgetting refers to the phenomenon where a model loses its ability to recall previous knowledge after learning new information [4]. In financial systems, this could be disastrous, as losing insights into past market behavior can lead to significant financial losses. Intransigence, on the other hand, is the model's resistance to learning new classes while attempting to retain old knowledge, which is detrimental in dynamic environments like finance, where new data streams constantly introduce unfamiliar variables.

Research in data privacy, especially with the advent of the General Data Protection Regulation (GDPR), has driven the need for efficient unlearning mechanisms. This need is exacerbated in graph data where the erasure of one node or edge could impact the global structure. Related work in data deletion, differential privacy, and adversarial robustness provides a foundation, but applying these techniques to graph data remains a challenge. Emerging work in graph learning, including Graph Neural Networks (GNNs), has only recently begun addressing privacy-preserving and unlearning mechanisms. Ref. [5,6] have provided a summary of the most relevant research on federated unlearning. Ref. [7] has proved that GNN is very successful in representing complex relationships in machine learning. When GNN framework is combined with treasury [8] and crypto trading [9], it becomes very powerful in machine unlearning.

Ref. [10–12] use unique methods called PROJECTOR and GRAPHEDITOR. In PROJECTOR [10], it uses projection techniques to remove specific nodes, ensuring no trace in the model parameters. In GRAPHEDITOR [11], it manages dynamic graphs and enables node/edge deletion and feature updates. The next major categories are the guaranteeing certified unlearning. The most famous is the CEU framework [13,14], which introduces a single-step update methodology for the removal of specific edges [5].

Machine unlearning is further exemplified by GUIDE [15], FedLU [16] and GNNDELETE [17]. Ref. [15] uses inductive graph unlearning in dynamic graphs. Ref. [16] features federated learning for knowledge-graph embedding for data heterogeneity. Ref. [17] introduces optimization strategies for node/edge deletions without loss of knowledge. The optimization strategies introduced in [17] is like the one used in self-supervised learning in [18]. Ref. [19] introduces Graph Scattering Transform which focus on mathematical robustness in unlearning, while [20] proposes Graph Influence Function that emphasizes influence function in unlearning. Ref. [21] proposes GraphEraser which stresses efficient partitioning mechanisms for unlearning in graph data.

### **Class Incremental Learning (CIL)**

Class Incremental Learning (CIL) refers to a paradigm in which a model learns new classes of data as they become available, without needing to retrain on the entire dataset from scratch. This approach is highly valuable in environments like finance, where the data landscape is fluid and non-stationary, with new financial instruments, market conditions, or types of fraud emerging over time. The ability to learn continuously, while preserving the knowledge of previously learned classes, is crucial for maintaining model performance in such dynamic environments.

One of the primary challenges in CIL is avoiding catastrophic forgetting, which occurs when the introduction of new classes causes the model to forget previously learned information. To mitigate this, several strategies have been developed. Regularization-based approaches constrain the updates to the model's parameters to prevent them from deviating too far from the values learned from previous classes. For example, Elastic Weight Consolidation (EWC) penalizes changes to parameters that are important for retaining knowledge of past classes, thereby allowing the model to learn new information without erasing the old. Another technique, learning without Forgetting (LwF), maintains knowledge of prior classes by combining the new data with the model's predictions from earlier data, ensuring that the model updates are balanced between old and new knowledge.

Memory replay is another strategy often employed in CIL. In this approach, a small subset of old data is stored and periodically replayed to the model during training on new data. This allows the model to refresh its memory of previous classes, reducing the likelihood of forgetting. Incremental Classifier and Representation Learning (iCaRL) is a prominent example of this method, as it integrates both memory replay and a classifier that incrementally updates the model's knowledge. Another notable approach is Gradient Episodic Memory (GEM), which stores gradients from previous tasks to ensure that the model's learning on new data does not interfere with past knowledge.

In some cases, models are designed to expand their architecture dynamically as new classes are introduced. Progressive neural networks, for instance, add new parameters to the model when a new class is learned. This approach allows the model to grow and adapt without overwriting older knowledge, thus preserving past learning. This strategy can be particularly useful in financial models where new assets or market segments may be introduced over time.

In the financial domain, CIL plays a critical role in applications such as portfolio management and fraud detection. In portfolio management, the introduction of new financial instruments like cryptocurrencies or tokenized assets requires machine learning models to learn the properties of these assets while retaining knowledge of traditional asset classes such as stocks and bonds. A portfolio management model that employs CIL can incrementally learn the dynamics of new asset classes without needing to retrain from scratch, making it highly efficient and adaptable to the constantly evolving market landscape.

Fraud detection is another area where CIL is invaluable. Financial institutions must continually adapt their fraud detection models to respond to new types of financial crimes and cyber threats. With CIL, these models can incorporate new fraud patterns over time while preserving knowledge of existing fraud cases. This ensures that the system remains robust and capable of detecting both old and new forms of fraud.

### **Machine Unlearning (MU)**

Machine Unlearning (MU) refers to the process by which a machine learning model forgets specific data points or the influence they exert on the model. MU has gained significant attention due to growing privacy concerns and regulatory requirements, which mandate that users should have the ability to request the removal of their data from the systems of financial services and other organizations. MU is crucial in ensuring that not only is the user's data deleted from the dataset, but its impact on the model is also erased, thus fully complying with privacy laws.

MU can be implemented using a variety of techniques, each with its own trade-offs between computational efficiency and accuracy. The most straightforward approach is exact unlearning, in which the model is retrained from scratch after the requested data points are removed. While this method is effective in completely erasing the influence of the data, it is computationally expensive and impractical for large datasets or models that require frequent updates, such as those used in financial applications.

To address the computational burden of exact unlearning, approximate unlearning methods have been developed. These methods seek to estimate and remove the influence of specific data points without fully retraining the model. One common approach involves storing information about the influence of each data point during the training process. When a data point needs to be unlearned, its influence can be subtracted from the model's parameters, resulting in a model that is approximately the same as if the data had never been included. This approach is computationally efficient and particularly suited for real-time financial models, where decisions need to be made quickly, and full retraining is not feasible.

Another technique used in MU involves randomization and sharding. In this method, the training data is divided into random subsets or shards, and each shard is used to train a separate model. When a data point needs to be unlearned, only the shard containing that data needs to be retrained or removed, leaving the other shards unaffected. This method reduces the computational cost of unlearning while ensuring that the model retains much of its original accuracy.

The financial industry stands to benefit significantly from the development of efficient MU methods, particularly in the areas of data privacy compliance and security. Financial institutions handle vast amounts of sensitive customer data, making compliance with privacy regulations like the GDPR and CCPA a top priority. MU provides a mechanism for ensuring that when a customer requests the deletion of their data, the machine learning models using that data also forget its influence. This is particularly important for institutions that use machine learning for credit scoring, risk assessment, or personalized financial recommendations, where customer data plays a central role.

MU is also valuable in mitigating the effects of data poisoning attacks, where adversaries intentionally insert malicious data into the training set to manipulate the model's predictions. In financial systems, where predictive models are used for trading, fraud detection, and investment recommendations, the impact of such attacks can be severe. MU allows the system to efficiently remove the poisoned data from the model, restoring its original performance and preventing future damage.

### **Synergies between Class Incremental Learning and Machine Unlearning**

While CIL and Machine Unlearning address different aspects of machine learning, their integration can lead to significant benefits. CIL enhances a model's adaptability by enabling it to learn new classes without forgetting old ones, making it more versatile in dynamic environments. Machine Unlearning, on the other hand, provides the flexibility to remove specific data points or classes from a model, ensuring data privacy and compliance with regulations. When combined, these two capabilities allow for the development of machine learning systems that can both learn and forget as needed.

One potential synergy is the use of CIL techniques to facilitate Machine Unlearning. For example, regularization-based methods used in CIL, which constrain the model's parameter updates, can also be applied to unlearning. By limiting how much the model changes when removing data, these methods can help ensure that the removal of specific data points does not significantly impact the model's performance on the remaining data. Additionally, rehearsal-based methods in CIL, which involve maintaining a subset of past data, can be adapted to support unlearning by allowing for the selective removal of specific data points from the rehearsal set.

Another synergy lies in the use of dynamic architectures, a common approach in CIL, to support unlearning. Dynamic architectures, which expand, or contract based on the model's needs, can facilitate the addition and removal of classes or data points. For instance, when a new class is added, the model can allocate additional resources to learn it, and when a class or data point needs to be forgotten, these resources can be reallocated or pruned, thus maintaining the model's efficiency.

Despite the potential synergies between CIL and Machine Unlearning, several challenges remain. One of the primary challenges in CIL is achieving a balance between learning new information and retaining old knowledge. Methods that overly emphasize retaining old knowledge can limit the model's ability to learn new classes effectively, while those that prioritize new classes can lead to significant forgetting. In the context of Machine Unlearning, ensuring complete and efficient unlearning is difficult. Current techniques may not fully eliminate the influence of specific data points, leading to potential privacy concerns.

Another challenge is the computational complexity associated with both CIL and Machine Unlearning. While these techniques aim to reduce the need for full retraining, they still require significant resources, particularly for large-scale models and datasets. Developing more efficient algorithms and architectures that can support both incremental learning and unlearning without excessive computational overhead is an important area of future research.

### **Conclusion**

The integration of CIL and Machine Unlearning has numerous potential applications across various domains. In autonomous systems, such as self-driving cars, the ability to learn new scenarios or objects incrementally while also forgetting outdated or irrelevant information can lead to more

adaptable and safer systems. In healthcare, CIL can help diagnostic models incorporate new medical knowledge as it becomes available, while Machine Unlearning can ensure that sensitive patient data can be removed from models when necessary.

Future research could focus on developing hybrid models that natively support both CIL and Machine Unlearning. Such models could dynamically adjust their structure and parameters to accommodate new information and remove unwanted data as needed. Additionally, exploring the use of meta-learning and lifelong learning techniques could enhance the adaptability and efficiency of these systems, enabling them to better handle the complexities of real-world data.

The synergies between Class Incremental Learning and Machine Unlearning hold significant potential for advancing the capabilities of machine learning systems. By combining the adaptability of CIL with the data privacy and security benefits of Machine Unlearning, it is possible to create more robust, flexible, and compliant models. While challenges remain in terms of balancing learning and forgetting, as well as managing computational complexity, ongoing research in this area promises to unlock new possibilities for the development of intelligent systems capable of both learning and forgetting as required.

## References

1. E. Belouadah, A. Popescu and I. Kanellos, "A comprehensive study of class incremental learning algorithms for visual tasks," *Neural Networks*, vol. 135, p. 38–54, 2021.
2. G. M. Van de Ven, T. Tuytelaars and A. S. Tolias, "Three types of incremental learning," *Nature Machine Intelligence*, vol. 4, p. 1185–1197, 2022.
3. D.-W. Zhou, Q.-W. Wang, Z.-H. Qi, H.-J. Ye, D.-C. Zhan and Z. Liu, "Deep class-incremental learning: A survey," *arXiv preprint arXiv:2302.03648*, 2023.
4. R. Aljundi, K. Kelchtermans and T. Tuytelaars, "Task-free continual learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.
5. N. Li, C. Zhou, Y. Gao, H. Chen, A. Fu, Z. Zhang and Y. Shui, "Machine Unlearning: Taxonomy, Metrics, Applications, Challenges, and Prospects," *arXiv preprint arXiv:2403.08254*, 2024.
6. T. Shaik, X. Tao, H. Xie, L. Li, X. Zhu and Q. Li, "Exploring the landscape of machine unlearning: A comprehensive survey and taxonomy," *arXiv preprint arXiv:2305.06360*, 2023.
7. Z. Wang, Y. Zhu, Z. Li, Z. Wang, H. Qin and X. Liu, "Graph neural network recommendation system for football formation," *Applied Science and Biotechnology Journal for Advanced Research*, vol. 3, p. 33–39, 2024. doi: 10.5281/zenodo.12198843
8. Z. Li, B. Wang and Y. Chen, "Incorporating economic indicators and market sentiment effect into US Treasury bond yield prediction with machine learning," *Journal of Infrastructure, Policy and Development*, vol. 8, p. 7671, 2024.
9. Z. Li, B. Wang and Y. Chen, "A Contrastive Deep Learning Approach to Cryptocurrency Portfolio with US Treasuries," *Journal of Computer Technology and Applied Mathematics*, vol. 1, pp. 1-10, 2024.
10. W. Cong and M. Mahdavi, "Efficiently forgetting what you have learned in graph representation learning via projection," in *International Conference on Artificial Intelligence and Statistics*, 2023.
11. W. Cong and M. Mahdavi, "Grapheditor: An efficient graph representation learning and unlearning approach".
12. Y. Wei, X. Gu, Z. Feng, Z. Li and M. Sun, "Feature Extraction and Model Optimization of Deep Learning in Stock Market Prediction," *Journal of Computer Technology and Software*, vol. 3, 2024.
13. E. Chien, C. Pan and O. Milenkovic, "Certified graph unlearning," *arXiv preprint arXiv:2206.09140*, 2022.
14. K. Wu, J. Shen, Y. Ning, T. Wang and W. H. Wang, "Certified edge unlearning for graph neural networks," in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2023.
15. X. Zhu, G. Li and W. Hu, "Heterogeneous federated knowledge graph embedding learning and unlearning," in *Proceedings of the ACM web conference 2023*, 2023.
16. C.-L. Wang, M. Huai and D. Wang, "Inductive graph unlearning," in *32nd USENIX Security Symposium (USENIX Security 23)*, 2023.
17. J. Cheng, G. Dasoulas, H. He, C. Agarwal and M. Zitnik, "Gnndelete: A general strategy for unlearning in graph neural networks," *arXiv preprint arXiv:2302.13406*, 2023.
18. H. Zhao, Y. Lou, Q. Xu, Z. Feng, Y. Wu, T. Huang, L. Tan and Z. Li, "Optimization Strategies for Self-Supervised Learning in the Use of Unlabeled Data," *Journal of Theory and Practice of Engineering Science*, vol. 4, p. 30–39, 2024.
19. C. Pan, E. Chien and O. Milenkovic, "Unlearning graph classifiers with limited data resources," in *Proceedings of the ACM Web Conference 2023*, 2023.

20. J. Wu, Y. Yang, Y. Qian, Y. Sui, X. Wang and X. He, "Gif: A general graph unlearning strategy via influence function," in Proceedings of the ACM Web Conference 2023, 2023.
21. M. Chen, Z. Zhang, T. Wang, M. Backes, M. Humbert and Y. Zhang, "Graph unlearning," in Proceedings of the 2022 ACM SIGSAC conference on computer and communications security, 2022.

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