

Retrain AI Systems Responsibly! Use Sustainable Concept Drift Adaptation Techniques

Lorena Poenaru-Olaru*, June Sallou*, Luis Cruz*, Jan S. Rellermeyer †, Arie van Deursen*

*Software Engineering, TU Delft, Netherlands

†Dependable and Scalable Software Systems, Leibniz University Hannover, Germany

L.Poenaru-Olaru@tudelft.nl, j.sallou@tudelft.nl, L.Cruz@tudelft.nl, rellermeyer@vss.uni-hannover.de, arie.vandeursen@tudelft.nl

Abstract—Deployed machine learning systems often suffer from accuracy degradation over time generated by constant data shifts, also known as concept drift. Therefore, these systems require regular maintenance, in which the machine learning model needs to be adapted to concept drift. The literature presents plenty of model adaptation techniques. The most common technique is periodically executing the whole training pipeline with all the data gathered until a particular point in time, yielding a massive energy footprint. In this paper, we propose a research path that uses concept drift detection and adaptation to enable sustainable AI systems.

Index Terms—sustainable model retraining, concept drift adaptation, sustainable model maintenance

I. INTRODUCTION

During the previous decade, there has been an exponential increase in machine learning adoption in industry [7]. Once deployed into production, machine learning systems suffer from vulnerabilities related to changes in data distribution over time [4]. These changes are known as *concept drift* [2] and they are influenced by external factors, such as geopolitical decisions, the macro-economical landscape, or user behavior changes. For instance, the outbreak of COVID-19 was the primary source of concept drift in data related to businesses within the hospitality sector¹.

The occurrence of concept drift is a tremendous threat to the performance of deployed machine learning models. The machine learning algorithms used to train the models assume that the distribution of the evaluation data is similar to the distribution of the data they have learned from during the training process. A violation of this assumption leads to a high probability of obtaining inaccurate predictions. Since machine learning models are dependent on these algorithms, they are prone to inaccurate predictions when evaluated on real-world data, where concept drift is ubiquitous.

To react to concept drift, machine learning models running in production need to be constantly adapted [2]. Previous studies [1]–[3] discuss concept drift adaptation from the **dataset adaptation perspective**, the **model adaptation perspective**, and the **frequency adaptation perspective**. The

aim of this paper is to analyze these adaptation techniques from a sustainability perspective and to highlight promising research possibilities on sustainable model adaptation.

II. CONCEPT DRIFT ADAPTATION TECHNIQUES

This section presents the state of the art in maintaining deployed machine learning models, focusing on the three aforementioned adaptation perspectives, namely the dataset, the model, and the frequency adaptation. We further analyze each technique from the sustainability perspective. We present our sustainability vision also in Figure 1.

A. Dataset Adaptation Perspective

a) *Definition*: The most commonly used dataset adaptation techniques are **periodically enlarging the training set**, the **gradual forgetting**, and the **abrupt forgetting** [3], [2].

The *periodically enlarging the training set* technique, presented in [3], implies a constant extension of the training dataset over time. The machine learning model needs to be periodically retrained on the expanded dataset.

The two forgetting techniques introduced in [2] work under the assumption that the most recent data is the most relevant to what the model will be evaluated on in the future. The *gradual forgetting* technique is similar to the previous one, but the difference between the two is that, although all the samples are stored over time, they are weighted when the model is retrained. The weight corresponds to the recency of their inclusion in the training dataset. Therefore, the most current samples receive the highest importance weight. When it comes to the *abrupt forgetting* technique, the older samples are completely discarded, and only the most recent samples are used to retrain the model. Thereby, with every iteration, the adaptation dataset size remains relatively constant.

b) *Sustainability Vision*: Among the dataset adaptation techniques, the most promising from a sustainability perspective is the *abrupt forgetting*. Verdecchia et al. [6] conducted an exploratory analysis to understand the effects of training dataset sizes on energy efficiency. The authors concluded that less energy is consumed when training on smaller datasets. Out of the three dataset adaptation techniques, abrupt forgetting is the only one to keep the dataset size relatively stable over time, since the model adaptation is done only on the most

¹Assessing the Effects of COVID-19 on Restaurant Business From Restaurant Owners' Perspective, Sardar, Sazu and Ray, Rudrendu and Hasan, Md Kamrul and Chitra, Shital Sohani and Parvez, ATM Shahed and Avi, Md Ashikur Rahman, Frontiers in Psychology

recent data. The *periodically enlarging the training set* and the *gradual forgetting* techniques imply that the training set size is constantly growing, which results in higher energy costs with every model adaptation iteration. Future research should focus on investigating which of the three techniques leads to the best prediction performance and, if the case, what is the trade-off between the accuracy loss and energy savings.

B. Model Adaptation Perspective

a) *Definition:* In terms of model adaptation perspective, a distinction can be made between **model restart** and **model retraining** [1]. Model restart implies that the old machine learning model is completely discarded, and a new model is built from scratch. This involves all the steps related to data processing and model exploration, such as different algorithms evaluation and model architecture experimentation. Model retraining implies that the existing model is retrained on the new training data. Therefore, through model retraining, the need to perform model exploration is eliminated.

b) *Sustainability Vision:* Retraining models has the potential to save more energy than restarting the models. Researchers at Facebook claim that the model exploration step consumes plenty of energy due to the concurrent investigation of multiple configurations [7]. Model exploration is a crucial process in the model restart technique. Therefore, restarting a model comes with high energy consumption. Model retraining eliminates the need for searching for the best algorithm since the existing algorithm is still employed, lowering the computational demand. Consequently, the energy consumption is reduced, making the model retraining a more sustainable technique to adapt to concept drift. Yet, it is still unclear which technique leads to better model performance. Thereby, we argue that more research is needed to understand which technique works better in terms of the model's performance and how much energy is saved when using model retraining instead of model restart.

C. Frequency Adaptation Perspective

a) *Definition:* When it comes to the frequency adaptation perspective, one can differentiate between **blind adaptation** and **informed adaptation** [2]. The blind adaptation implies that the model is adapted on a periodic basis, for instance, every month. The informed adaptation performs adaptation only when necessary. The necessity of adapting the model is signaled by a monitoring tool called a concept drift detector. In [5], the authors present two different categories of concept drift detectors. The idea behind this technique is that the model should only be adapted when necessary, when concept drift is detected, instead of regularly.

b) *Sustainability Vision:* The model adaptation should be done only when needed (*informed adaptation*), not periodically (*blind adaptation*). Although popular among industry practitioners [4], the blind adaptation technique implies that the model should be adapted periodically without investigating the need for adaptation. The *informed adaptation* technique, on the contrary, performs drift adaptation only after concept

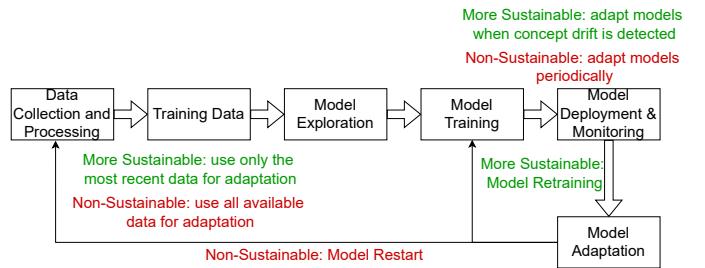


Fig. 1. Sustainable vs Non-Sustainable Concept Drift Adaptation Techniques

drift is detected, eliminating unnecessary model adaptation that consumes energy. However, the informed adaptation technique requires an additional step in the adaptation pipeline, namely the concept drift detector. Therefore, future research directions are threefold: investigating the most sustainable concept drift detector/s category from an energy consumption perspective, verifying the detector's ability to capture drift, and investigating the most sustainable frequency adaptation technique.

III. FUTURE STEPS

In this paper, we bring a sustainability perspective to the way we maintain deployed machine learning models against concept drift. We presented various methods used in the literature to update machine learning models running in production and evaluated them from three different perspectives, namely the dataset, the model, and the frequency adaptation perspectives. We analyzed these techniques from a sustainability point of view in terms of the energy consumption required to perform each adaptation technique. Further research should focus on studying the energy consumption for each of the aforementioned adaptation techniques. We believe the research community has overlooked the major impact on energy consumption that stems from keeping machine learning systems up to date. Future research should investigate which drift adaptation techniques are more sustainable while addressing performance trade-offs.

REFERENCES

- [1] Celik, B., Vanschoren, J.: Adaptation strategies for automated machine learning on evolving data. *IEEE transactions on pattern analysis and machine intelligence* (2021)
- [2] Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., Bouchachia, A.: A survey on concept drift adaptation. *ACM Computing Surveys (CSUR)* 46, 1 – 37 (2014)
- [3] Kavikondala, A., Muppalla, V., Prakasha, D.K., Acharya, V.: Automated retraining of machine learning models. *International Journal of Innovative Technology and Exploring Engineering* 8, 445–452 (10 2019)
- [4] Lyu, Y., Li, H., Sayagh, M., Jiang, Z.M.J., Hassan, A.E.: An empirical study of the impact of data splitting decisions on the performance of aiops solutions. *ACM Trans. Softw. Eng. Methodol.* 30 (2021)
- [5] Poenaru-Olaru, L., Cruz, L., van Deursen, A., Rellermeyer, J.S.: Are concept drift detectors reliable alarming systems? - a comparative study. In: *7th Workshop on Real-time Stream Analytics, Stream Mining, CER/CEP & Stream Data Management in Big Data* (2022)
- [6] Verdecchia, R., Cruz, L., Sallou, J., Lin, M., Wickenden, J., Hotellier, E.: Data-centric green ai an exploratory empirical study. In: *2022 International Conference on ICT for Sustainability (ICT4S)*. pp. 35–45 (2022)
- [7] Wu, C.J., et al.: Sustainable ai: Environmental implications, challenges and opportunities. In: *Proceedings of Machine Learning and Systems*. vol. 4, pp. 795–813 (2022)