Revisiting Replay and Gradient Alignment For Continual Pretraining of Large Language Models

Large language models (LLMs) require continual updates to remain relevant, yet the dominant paradigm of retraining from scratch is prohibitively costly in compute and energy. This work investigates continual pre-training (CPT), where models are incrementally updated on massive new datasets without discarding prior knowledge. The central challenge in CPT is the stability–plasticity dilemma: improving performance on new data while retaining competence on earlier tasks, thus avoiding catastrophic forgetting. The paper revisits two prominent strategies from the continual learning literature : experience replay and gradient alignment , and provides the first comprehensive study of their effectiveness at LLM scale.

In this paper, we evaluate sequential pre-training of Llama-based Spectra models ranging from 99M to 6B parameters across multilingual corpora (English, French, German, Arabic, Japanese; 100B tokens each). Experience replay is implemented using a scalable disk-backed buffer, enabling efficient interleaving of old and new examples across more than 500B tokens. To complement replay, we introduce an efficient implementation of Meta-Experience Replay (MER), which integrates Reptile-style meta-updates to encourage positive gradient alignment with negligible overhead. This combination aligns gradients across tasks, reducing interference and enhancing transfer.

Empirical analyses demonstrate several key findings. First, replay significantly mitigates forgetting: a 560M model with 25% replay achieves comparable stability to a 1B model trained without replay, showing replay can be a more compute-efficient investment than scaling parameters. Second, gradient alignment and replay are synergistic: MER consistently outperforms either method alone, achieving lower forgetting scores and better downstream generalization (e.g., HellaSwag, PiQA, PubMedQA). At 6B scale, MER with 50% replay surpasses joint training baselines, suggesting replay plus alignment can act as an implicit curriculum. Third, scaling studies reveal that moderate replay rates (25%) combined with gradient alignment yield the best stability–plasticity trade-off; investing further compute in replay is less effective than increasing model size. Importantly, gradient alignment adds nearly "free" improvements, as its overhead is negligible relative to training costs.

The work makes three contributions: (1) the first large-scale validation of gradient alignment in LLM pre-training, (2) an efficient MER implementation compatible with Megatron/NeoX infrastructure, and (3) a detailed compute-aware scaling analysis of replay versus model growth. Together, the findings establish replay as a cost-effective strategy for CPT and highlight gradient alignment as a lightweight yet powerful complement. Beyond advancing methodology, this study underscores the potential of CPT to reduce retraining costs and carbon footprint in maintaining foundation models. By demonstrating that LLMs can be continually pre-trained with stability, plasticity, and generalization, the paper paves the way for more sustainable, adaptive, and efficient large-scale language modeling.