ML Lab Research Proposal

1. Problem Statement
   Optimizing data mixtures for large language models remains computationally prohibitive because identifying the best mixture ratios typically requires repeated full-scale training runs guided by heuristic tuning. Existing methods are primarily optimized for language modeling loss such as perplexity, but reductions in perplexity do not reliably translate to improved downstream reasoning, coding, and mathematical performance. This project investigates whether parameter-space interpolation can serve as a high-fidelity proxy for mixed-data optimization by exploiting the shared first-order optimization dynamics between independently trained domain experts and mixed-domain training trajectories.
2. Proposed Approach
   We propose a framework that approximates mixed-data training using linear interpolation between independently trained domain-specific expert models. Given interpolation coefficients $\alpha \in \Delta K-1$, the merged parameters are constructed as $\Theta merge(\alpha) = \Theta base + \Sigma k\ \alpha k(\Theta k - \Theta base)$, allowing efficient exploration of the mixture simplex without retraining full models. The method is motivated by a second-order Taylor expansion analysis showing that mixed-data optimization and parameter interpolation share identical first-order gradient terms while discrepancies are confined to second-order curvature interactions involving Hessian-gradient products Hkgj. We additionally fit LightGBM regressors to model the performance surface across downstream benchmark domains and optimize a utility function over predicted benchmark scores.
3. Dataset
   We use a large-scale mid-training corpus partitioned into four primary domains: mathematics, code, supervised fine-tuning data, and web/other data. The benchmark suite spans reasoning, mathematical problem solving, coding, language understanding, and professional knowledge tasks, including GSM8K, MATH, HumanEval, MBPP, GPQA, MMLU, ARC, BBH, and TriviaQA. Data preprocessing includes domain clustering, token filtering, normalization, and hierarchical decomposition into fine-grained sub-clusters such as formal mathematics, synthetic study notes, code repositories, long chain-of-thought supervision, and academic papers.
4. Evaluation Metrics
   Performance is evaluated using average benchmark accuracy across five capability domains: math, code, knowledge, language, and reasoning. To validate the interpolation proxy, we measure rank consistency between predicted interpolation performance and actual full-training performance using Spearman rank correlation coefficients, targeting correlations above 0.9. Additional evaluation includes capability-specific benchmark scores, cross-domain transfer behavior, and analysis of Hessian interaction matrices and task-vector cosine similarities to quantify cross-domain interference and optimization orthogonality.
5. Compute Requirements
   The project requires approximately 2× AMD MI300X GPU accelerators with a combined memory budget near 384GB HBM memory to support distributed sparse Mixture-of-Experts mid-training experiments. We estimate approximately 10–14 days of

experimentation, including multiple 5B–25B token expert training runs, interpolation sweeps across the mixture simplex, downstream benchmark evaluation, and regression surface fitting. Dependencies include PyTorch, DeepSpeed, FlashAttention, distributed checkpoint interpolation utilities, Hessian-vector product implementations, and LightGBM-based regression models for capability prediction.

6. Expected Deliverables
   We expect to deliver a working framework for efficient data mixture optimization using parameter interpolation, trained domain-specialized expert checkpoints, and a final optimized mid-trained language model. We additionally plan to provide benchmark evaluation reports, visualization of capability landscapes projected over the mixture simplex, and empirical analysis of first-order versus second-order optimization behavior during interpolation-based search. The final output will include reproducible scripts for interpolation search, utility optimization, and downstream benchmark evaluation.
7. Risk & Mitigation
   The primary risk is that higher-order curvature interactions may cause parameter interpolation to diverge from the behavior of real mixed-data training, especially when optimization trajectories move into separate loss basins. Another risk is that domain-specific task vectors may not remain sufficiently orthogonal, reducing the reliability of interpolation-based ranking predictions. To mitigate these issues, we constrain expert training to short horizons with constant learning rates, preserve shared initialization across all experts, and empirically validate the approximation using Hessian interaction analysis, cosine similarity measurements, and full-training verification runs on sampled mixture configurations.