Semi-supervised Fine-tuning for Large Language Models

Junyu Luo[♡], **Xiao Luo**[♠ †], **Xiusi Chen**[◇], **Zhiping Xiao**[♣ †], **Wei Ju**[♡], **Ming Zhang**[♡ †]
[♡] State Key Laboratory for Multimedia Information Processing,
School of Computer Science, PKU-Anker LLM Lab, Peking University
[♠] University of California, Los Angeles [◇] University of Illinois Urbana-Champaign
[♣] University of Washington
Github Repository: `https://github.com/luo-junyu/SemiEvol`.

Abstract

Supervised fine-tuning (SFT) is crucial in adapting large language models (LLMs) to a specific domain or task. However, only a limited amount of labeled data is available in practical applications, which poses a severe challenge for SFT in yielding satisfactory results. Therefore, a data-efficient framework that can fully exploit labeled and unlabeled data for LLM fine-tuning is highly anticipated. Towards this end, we introduce a semi-supervised fine-tuning (SemiFT) task and a framework named SEMIEVOL for LLM alignment from a *propagate-and-select* manner. For knowledge propagation, SEMIEVOL adopts a bi-level approach, propagating knowledge from labeled data to unlabeled data through both in-weight and in-context methods. For knowledge selection, SEMIEVOL incorporates a collaborative learning mechanism, selecting higher-quality *pseudo-response* samples. We conducted experiments using GPT-4o-mini and Llama-3.1 on seven general or domain-specific datasets, demonstrating significant improvements in model performance on target data. Furthermore, we compared SEMIEVOL with SFT and self-evolution methods, highlighting its practicality in hybrid data scenarios.

1 Introduction

Supervised fine-tuning (SFT) is a crucial method for enhancing large language models' (LLMs) performance on instructional or domain-specific tasks (Raffel et al., 2020; Chung et al., 2024), playing a vital role in adapting LLMs for specific scenarios. However, SFT relies on a substantial amount of annotated labeled data, which can be increasingly costly in real-world applications (Honovich et al., 2023; Kung et al., 2023). While existing LLMs often employ unsupervised pretraining methods (Devlin, 2018; Radford et al., 2019; Brown, 2020) to improve their capabilities, this approach typically requires vast datasets and substantial computational resources, making it impractical for scenarios with limited accessible samples.

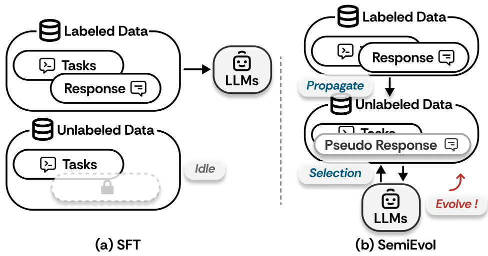

Figure 1: Comparison of SEMIEVOL with previous SFT methods. SEMIEVOL enables interaction between diverse data types for superior performance evolution.

In practice, however, it often presents a hybrid situation, where a small amount of labeled data coexists with a relatively larger volume of unlabeled data. On the one hand, when deploying LLMs to new target tasks, a limited amount of task-specific annotations can be valuable without incurring excessive costs (Perlitz et al., 2023; Kung et al., 2023). On the other hand, during the continuous inference process of LLMs, a substantial amount of unlabeled data accumulates (Tao et al., 2024; Honovich et al., 2023; Wang et al., 2023). Effectively leveraging the labeled data to enhance model performance on unlabeled data, while simultaneously selecting high-quality unlabeled samples, can improve LLMs' performance in target scenarios, offering substantial practical utility. Therefore, we aim to address the following question:

> *Can LLMs evolve in a real-world scenario of limited labeled data and abundant unlabeled data?*

Designing an evolution framework for hybrid-data scenarios is non-trivial due to the following reasons: First, semi-supervised learning (Kipf and Welling, 2016; Shi et al., 2023), which has been widely studied in machine learning, primarily fo-

† Corresponding authors.