STAR: Constraint LoRA with Dynamic Active Learning for Data-Efficient Fine-Tuning of Large Language Models

ACL 2024 Bangkok, Thailand

SOUTHEAST UNIVERSITY

Linhai Zhang♠♢* Jialong Wu♠♢* Deyu Zhou♠♢(✉) Guoqiang Xu♡

♠School of Computer Science and Engineering, Southeast University, Nanjing, China
♢Key Laboratory of New Generation Artificial Intelligence Technology and Its Interdisciplinary Applications (Southeast University), Ministry of Education, China
♡SANY Group Co., Ltd.

Introduction

Though Large Language Models (LLMs) have demonstrated the powerful capabilities of few-shot learning through prompting methods, supervised training is still necessary for complex reasoning tasks. Because of their **extensive parameters** and **memory consumption**, both Parameter-Efficient Fine-Tuning methods and Memory-Efficient Fine-Tuning methods have been proposed for LLMs. Nevertheless, the issue of large annotated data consumption, the aim of Data-Efficient Fine-Tuning, remains unexplored.

One obvious way is to combine the PEFT method with active learning. However, as shown in Figure 1, the experimental results show that such a combination is **not trivial and yields inferior results**.

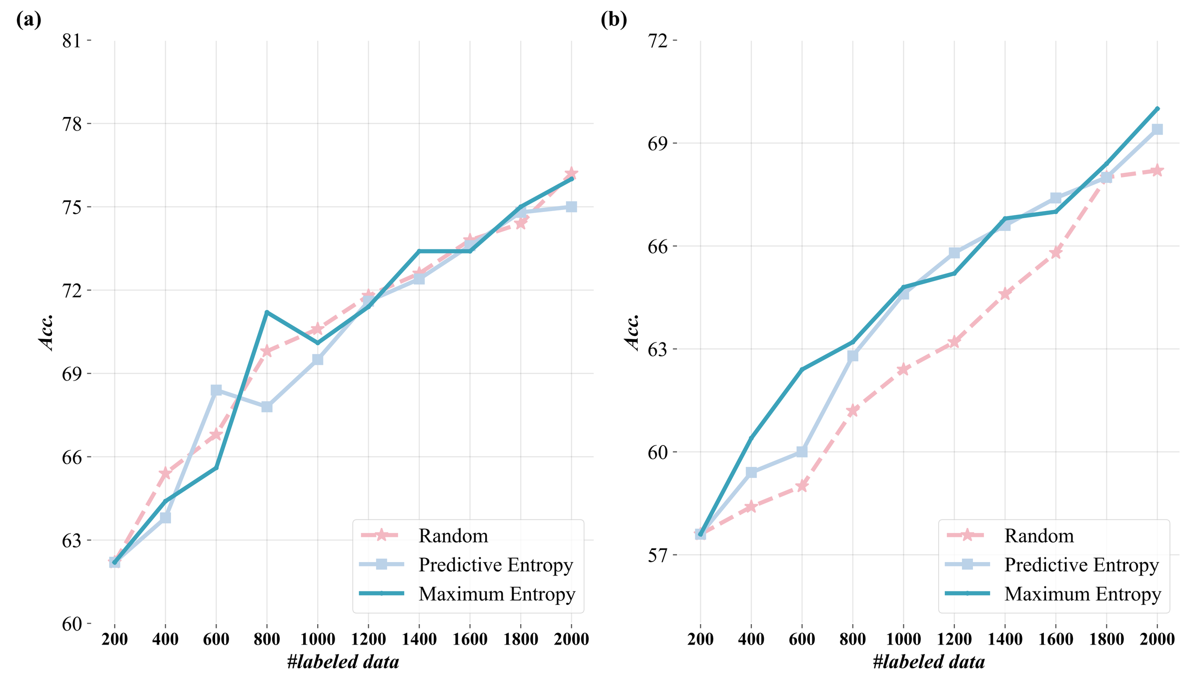

Figure 1. (a) Active learning combined with LoRA compared to passive learning. (b) Active learning combined with full parameter tuning compared to passive learning.

Through probe experiments, we hypothesize the observations can be explained by two main reasons: Uncertainty Gap and Poor Model Calibration.

To address the aforementioned issues, we propose con**S**train**T** LoRA with dynamic **A**ctive lea**R**ning (**STAR**), a novel approach for effectively integrating uncertainty-based active learning and LoRA.

Probing PEFT on Prediction Uncertainty

As uncertainty-based AL methods mainly depend on the confidence or uncertainty of model predictions to select examples during each iteration, it is straightforward to probe the confidence and uncertainty of model predictions during AL iterations.

Probing with Prediction Confidence. As shown in Figure 2, the first probe experiment is designed to explore whether the model prediction confidence of the *PEFT* method exhibits issues compared to *Few-shot* methods. The prediction confidence CF is measured by the maximum between the output probability on token "*true*" and "*false*".

$$CF = max(p_{true}, p_{false}) \tag{1}$$

where p_{true} and p_{false} denotes the probabilities of token "*true*" and "*false*", respectively.

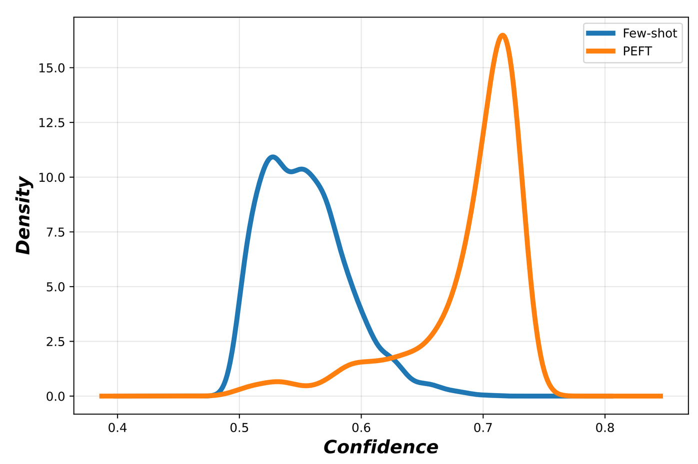

Figure 2. Density plot of confidence for wrong predictions.

The *PEFT* method is overconfident compared to the *Few-shot* method, where the confidence of the wrong prediction is as high as 70%, which indicates a ***model calibration issue***.

Probing with Prediction Entropy. As shown in Figure 3, the second probe experiment is designed to investigate the change of prediction entropy of *PEFT* model during active learning iteration. The Maximum Entropy(ME) is employed as the uncertainty during active learning. Nine rounds of iteration are performed with 500 examples selected during each iteration. The *PEFT* model is trained with 500 examples at the beginning as a warm-up.

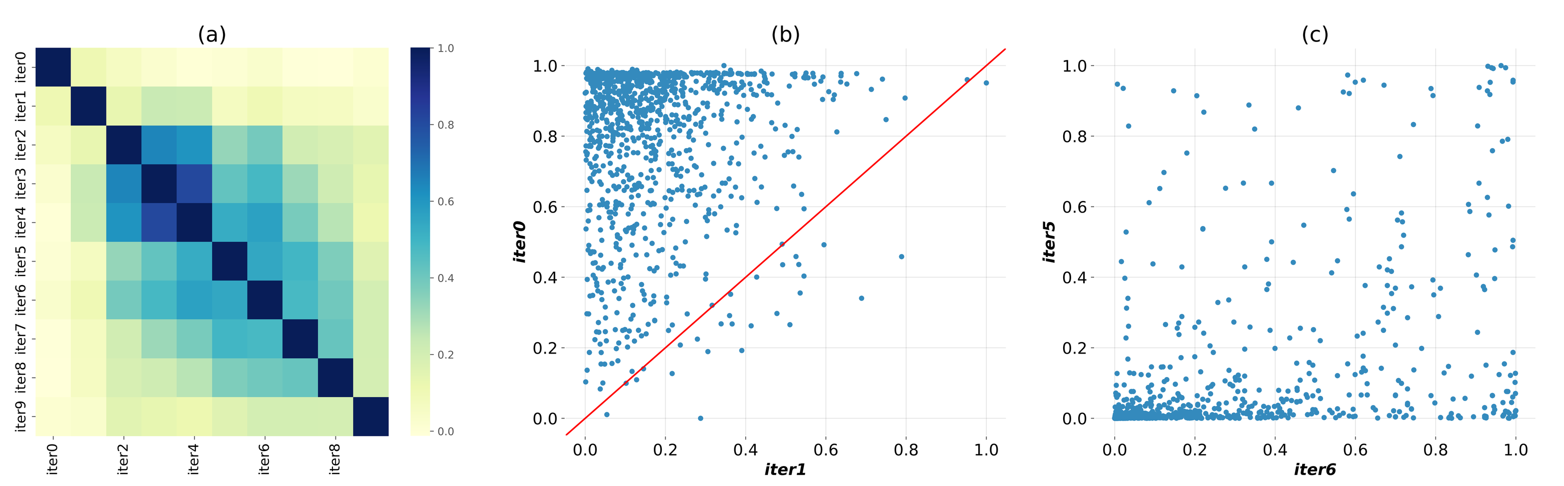

Figure 3. (a) Heatmap of correlation between prediction entropy across different iterations; (b) Scatter plot for prediction entropy between base model (Iter0) and model after first iteration (Iter1); (c) Same as (b), except values are taken from Iter5 and Iter6.

As we can observe in Figure 3 (a), the correlation between the base model (model without PEFT tuning) and models after AL iteration is close to 0, which indicates ***a clear gap between the base model and PEFT model***. This phenomenon is even clear with the scatter plot in Figure 3, where the dots in Figure 3 (b) should appear around the red line but appear in the upper triangular region. In Figure 3 (c), the correlation coefficients of entropy between the two iterations become relatively normal, which is consistent with Figure 3 (a), suggesting that the gap between iterations has been alleviated.

The framework of STAR

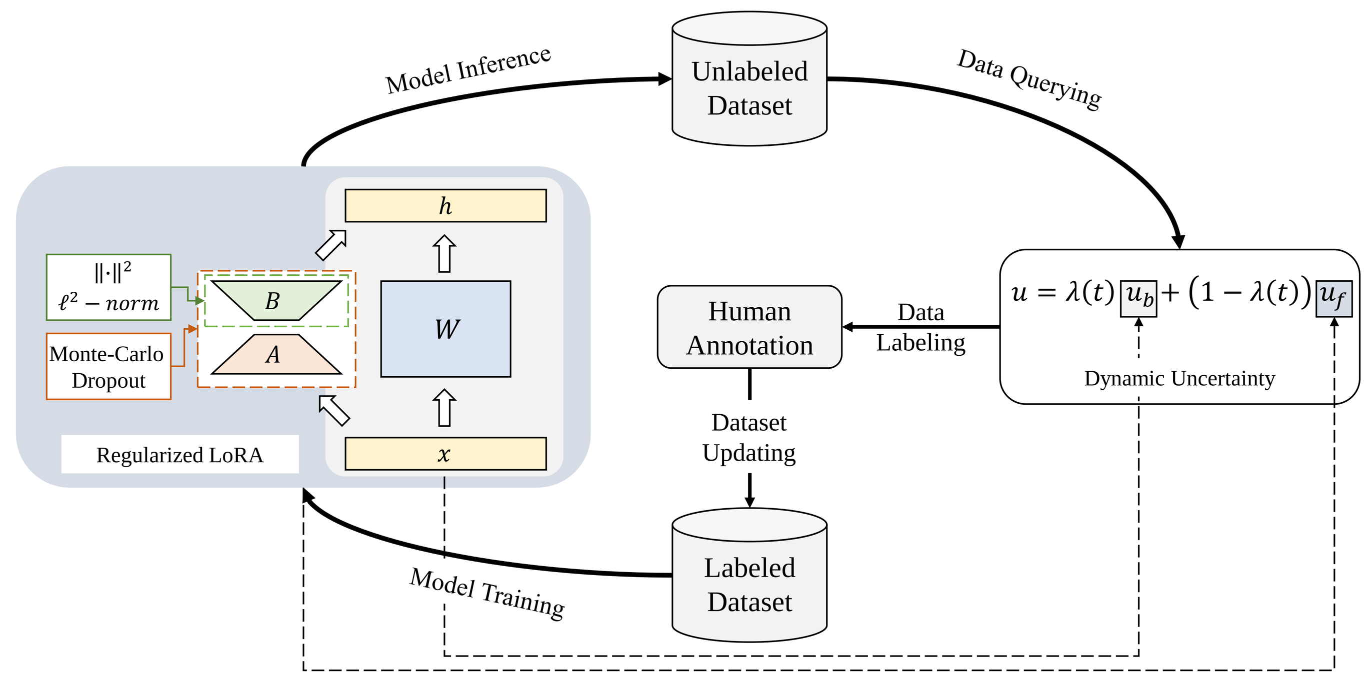

Figure 4. The framework of the proposed method.

Dynamic Uncertainty Measurement.To address the issue of ***uncertainty gap***, we proposed a dynamic uncertainty measurement to integrate the uncertainty of the frozen LLM (base model) and the uncertainty of LLM fine-tuned with LoRA (full model) dynamically based on the AL iteration. As the iteration of active learning increases, the uncertainty of the full model becomes more reliable, which is similar to the zero-initialized attention weight in LLaMA adapter.

$$\mu = \lambda(t)\mu_b + (1-\lambda(t))\mu_f \tag{2}$$

where μ_b and μ_f denote the prediction uncertainty of the base model and the full model respectively. $\lambda(t) \in [0, 1]$ is a monotone decreasing function of AL iteration t.

Calibration with Hybrid Regularization. To address the issue of ***poor model calibration***, we propose a hybrid regularization method during PEFT training. The PEFT model demonstrates a pronounced tendency toward over-confidence, which indicates that the model is over-fitting.

For the B matrix, which is zero-initialized, a L^2 norm weight decay is employed.

$$B_t \leftarrow B_{t-1} - \gamma(g_{t-1} - \beta B_{t-1}) \tag{3}$$

where g_{t-1} denotes the normalized gradient acquired from the standard Adam optimizer, and β denotes the strength of regularization.

For the A matrix, which is randomly Gaussian initialized $N(0, 1)$, the Monte-Carlo dropout (MC dropout) is adopted for more robust uncertainty estimation.

$$\begin{aligned} \mu_f &= \frac{1}{K}\sum_k \mu_f^{(k)} \\ \mu_f^{(k)} &= \mathit{ME}(\mathrm{LLM}(x|\hat{A}_k, \hat{B}_k)) \end{aligned} \tag{4}$$

where K denotes the number of feedforward propagations during the inference stage, $\mu_f^{(k)}$ denotes the uncertainty estimated at k-th feedforward, $\hat{A}_k$ and $\hat{B}_k$ denote the LoRA matrices sampled from A and B with dropout unit activated.

Experimental Result

Table 1 presents a detailed comparison of different methods' performance, evaluated across three different datasets: GSM8K, BoolQ, and OpenBookQA.

Method	GSM8K		BoolQ		OpenBookQA	
	AUC	RIPL	AUC	RIPL	AUC	RIPL
Random	27.37	-	60.46	-	63.44	-
Predictive Entropy	27.30	-0.09	58.39	-5.24	63.05	-1.07
w/ STAR	<u>28.40</u>	<u>1.42</u>	<u>61.84</u>	<u>3.49</u>	<u>64.86</u>	<u>3.88</u>
Maximum Entropy	27.16	-0.28	60.65	0.48	63.36	-0.22
w/ STAR	**28.83**	**2.01**	**61.91**	**3.67**	**66.17**	**7.47**

Table 1. The performance of different methods in a passive learning setup in terms of the AUC and RIPL. The optimal results among all methods are **bolded** and the second-best results are <u>underlined</u>.

Figure 5 shows the learning curves for corresponding AL methods on GSM8K, BoolQ, and OpenBookQA datasets, respectively.

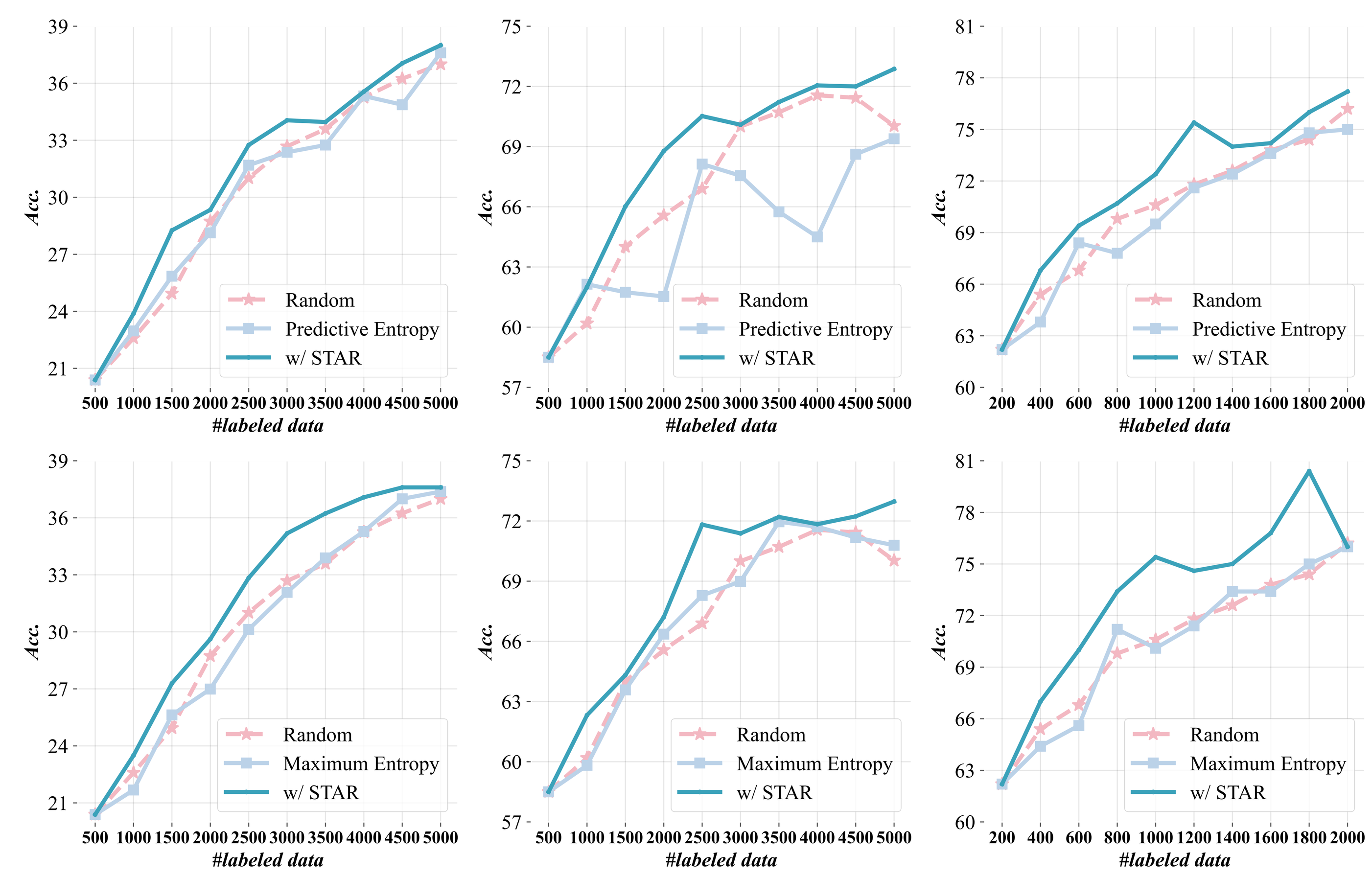

Figure 5. The Learning curves comparing the Predictive Entropy and Maximum Entropy methods, and each w/ STAR, against the RANDOM baseline. The first column corresponds to the GSM8K dataset, the second column to the BoolQ dataset, and the third column to the OpenBoolQA dataset.