

What's on your mind?

Choose an option:

- ☐ Lecture Review
- ☒ Ask Questions
- ☐ Test Yourself

TA-Lite: Academic Assistant

Logged in as: rarao@iu.edu

? Ask Questions

Ask any question about the course content:



What is the primary advantage of LoRA, and how does LoRA modify the behavior of pre-trained models dur



Answer

Great question! Let's dive into the primary advantage of LoRA and how it modifies the behavior of pre-train

First, consider this: The primary advantage of LoRA is its efficiency in reducing computational resources an because LoRA only updates a small subset of the model's parameters, making the fine-tuning process more

Now, let's talk about how LoRA modifies the behavior of the pre-trained models. LoRA essentially adds a new layers of the model. This new matrix acts like a filter that modifies the behavior of the model to better suit t

To understand this better, let's look at the lecture slides on page 55. The slides mention that LoRA's low-rar onto a subspace, which helps the model focus its attention on the relevant features for the new task.

So, my question to you is: How does the use of a low-rank matrix help in reducing the computational cost a fine-tuning? Take a moment to think about this and then share your thoughts.**Sources:**LLMs-06-fine-tur LLMs-06-fine-tuning.pdf, transcript-06-pretraining.docx, transcript-08-prompting.docx, transcript-06-pretr can refer to: YouTube from StatQuest:

- Transformer Neural Networks, ChatGPT's foundation, Clearly Explained!!! <https://www.youtube.com/y3Blue1Brown>:
- The scale of training LLMs <https://www.youtube.com/watch?v=qSn8dVI9V5I>
- Transformers (how LLMs work) explained visually | DL5 <https://www.youtube.com/watch?v=wjZofJX0v>
- Fine Tuning LLM Models – Generative AI Course <https://www.youtube.com/watch?v=iOdFUjiB0Zc>
- Stable Diffusion Crash Course for Beginners https://www.youtube.com/watch?v=dMkiOex_cKU

Medium Articles:

- Understanding LoRA — Low Rank Adaptation For ... <https://medium.com/data-science/understanding-finetuning-large-models-936bce1a07c6>
- LoRA: Low Rank Adaptation of Large Language Models <https://medium.com/@tayyibgondal2003/loral-language-models-33f9d9d48984>

Research Papers:

- KD-LoRA: A Hybrid Approach to Efficient Fine-Tuning with LoRA and Knowledge Distillation <http://arxiv>
- LoRA-drop: Efficient LoRA Parameter Pruning based on Output Evaluation <http://arxiv.org/abs/2402.07>



Previous Questions & Answers

Q: What is the primary advantage of LoRA, and how does LoRA modify the behavior of pre-trained models d... (10:37 PM)

What is up?