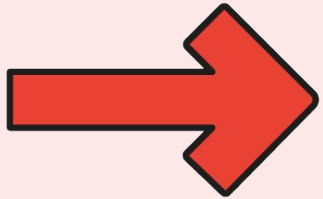
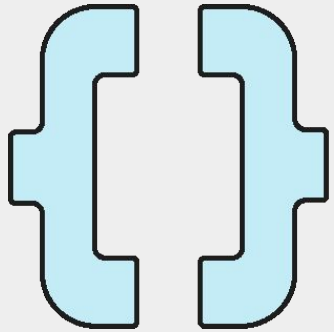
Google Developer Group
Chandigarh University

Gemini Study Jams

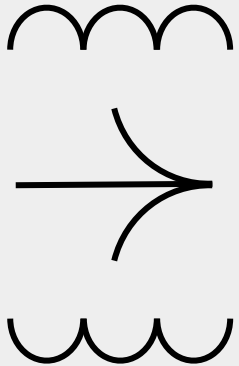
Session 3 - Supervised PEFT using LoRA Continued



What will we do today?

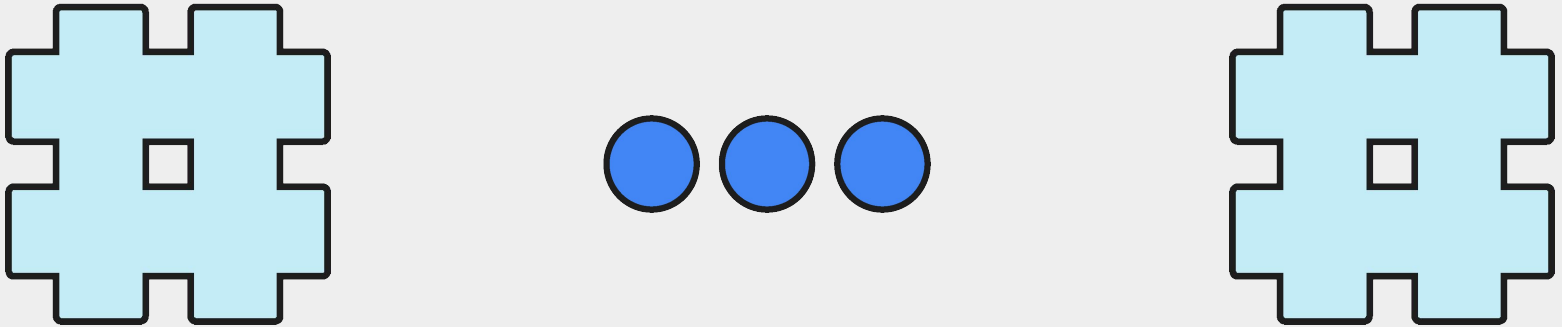


1. Cover theory of how LoRA works.
2. Look at LoRA hyperparameters.
3. Go through the first lab of the course.
4. Talk about problems that can hinder fine tuning.



How does LoRA work?

Let’s head to OneNote for an explanation …



LoRA Hyperparameters

1. **Rank r**

   r is the rank of the low-rank update

   $\Delta W = BA$

   Small r → fewer parameters, less expressive but cheaper

   Large r → more expressive, more parameters, but closer to full fine-tuning

   Usually r = 4, 8, 16 for large models

LoRA Hyperparameters

2. **Alpha (α) — scaling factor**

Controls the **magnitude of the LoRA adjustment**

$$W' = W + (\alpha / r)\, BA$$

Higher α → LoRA tweaks have **more influence**

Lower α → LoRA tweaks are subtle

LoRA Hyperparameters

3. Target modules

Which weight matrices LoRA is applied to:

- Usually **Q and V projections** in attention
- Sometimes also **FFN layers**
- Often skip K (less need to adapt keys)

Concerns with Fine Tuning

1. **High computational cost**

   Large models have billions of parameters (GPT-3: 175B, LLaMA 65B, etc.)

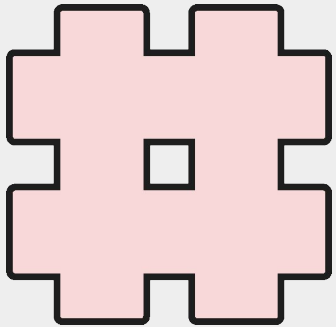
   Full fine-tuning updates every weight, which requires:

   a. Lots of GPU memory (storing activations and gradients)

   b. Long training times

   This often makes fine-tuning **impractical for most researchers or small teams**

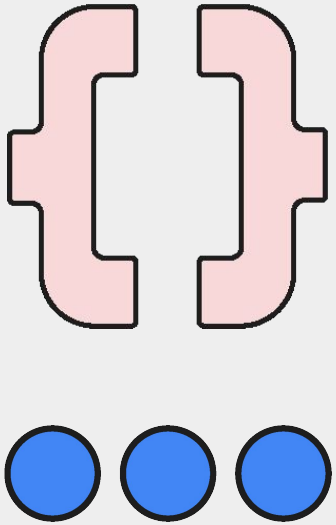
   Effect: You may need multiple GPUs, huge VRAM, or gradient checkpointing hacks.



Concerns with Fine Tuning

2. Risk of overfitting

Large models are very expressive; full fine-tuning on **small datasets** can easily overfit.



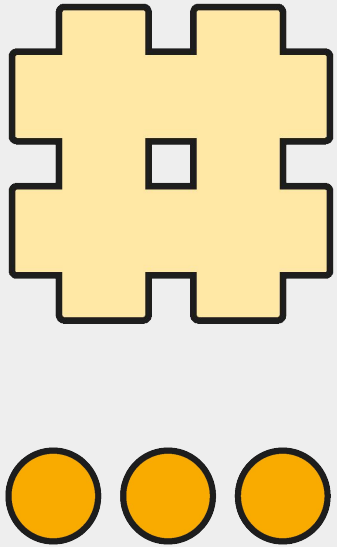
Fine-tuned model may perform **worse on general language tasks** or **forget knowledge** from pretraining.

Concerns with Fine Tuning

Catastrophic forgetting

Updating all weights can make the model “**forget**” what it learned during pre training.

The model may lose general capabilities (grammar, reasoning, factual knowledge) while adapting to a new task.



Code

Quiz

Doubts?

Ask away!

Your Turn!

Google Developer Group
Chandigarh University

Thank You!

Fine Tuning Basics & Supervised Fine Tuning - End