# LoRA for Supervised Learning

By Alisha Rath Presenting to Prof. Dr Chris Pollett

#### What is LoRA?







LoRA (Low-Rank Adaptation) is a method for fine-tuning large pre-trained models with a focus on reducing memory and computational costs. It introduces trainable low-rank matrices into the model's architecture while freezing the original pre-trained weights. Used for supervised learning tasks where labeled data is available.

## Benefits of LoRA

Low Computational Cost: Reduces the number of parameters to train.

**Efficient Fine-tuning:** Works well with large models like GPT, BERT, and LLaMA.

**Memory Efficiency:** Decreases the memory footprint of model updates, making it easier to use on resource-limited devices.

**Better Performance:** Preserves the original knowledge of the model while adapting to new tasks.

#### Mechanics of LoRA



where WWW is the original weight matrix.



Source: Arxiv





BΑ Δw (dx)(rxk) dxk

$$assume d = 3; k = 3$$





## **Application in Supervised Learning**





	=

Supervised Learning Basics: Learning a function from input-output pairs using labeled data.

LoRA's Role: Enables fine-tuning of large language models (LLMs) on specific datasets for classification, regression, or NLP tasks with less computational overhead.

**Example:** Fine-tuning a model for legal document classification using LoRA.

#### Steps in LoRA-based Supervised Learning



### Why LoRA is Superior



TUNING.

SPECIALIZED HARDWARE.

ADAPT A GENERAL-PURPOSE MODEL TO NEW SPECIFIC TASKS.

### LoRA in Natural Language Processing

**Text Classification:** Finetune models for text classification (e.g., sentiment analysis, legal case classification).

Summarization: Improve document summarization tasks by fine-tuning on task-specific datasets. Question Answering: Adapt LLMs for domainspecific question-answering with low memory and computational requirements.

### Challenges with LoRA







MODEL COMPLEXITY: WHILE LORA IS EFFICIENT, FINE-TUNING LARGE MODELS STILL REQUIRES SIGNIFICANT RESOURCES. SPECIALIZED DATASETS: PERFORMANCE HEAVILY RELIES ON THE QUALITY AND SIZE OF THE LABELED DATASET USED. **EVALUATION:** ENSURING ACCURATE EVALUATION METRICS FOR NEW TASKS IS CRUCIAL.









LoRA offers an efficient, scalable solution for fine-tuning large models on specific supervised tasks.

It reduces memory and computational costs without sacrificing model performance. Particularly useful for resourcelimited scenarios in NLP and other machine learning domains.



## Thank you