The $50,000 Question Has a $500 Answer
Model Adaptation: Cost vs Performance

Performance (Accuracy %) *(y-axis)*
Cost (USD) *(x-axis)*

RLHF

Full Fine-tuning

LoRA Fine-tuning

Sweet Spot:
0.1% cost, 98% performance

Prompt Engineering

Few-shot Prompting

Zero-shot Prompting

Key Insight: LoRA achieves 93% accuracy (98% of full FT) at 1% of the cost