

# LLM Inference Performance Comparison (LLaMA-2-13B)

