

Complete Results Summary

Metric	Base Model	Transformers+PEFT	Unsloth	vLLM	Ollama
Training Time (min)	-	53.3	43.2	-	-
Training Memory (GB)	-	2.24	1.44	-	-
Final Loss	-	1.413	1.367	-	-
Inference Latency (s)	-	1.501	0.918	0.246	0.440
Throughput (tok/s)	-	95.5	135.8	345.8	243.9
BLEU Score	4.32	5.23	6.59	-	-
ROUGE-L	0.169	0.195	0.236	-	-
BERTScore F1	0.8514	0.8209	0.8316	-	-