

HumanEval: Pass rate vs Avg latency (s per task)

