

HarmBench LLM Judge - Attack Success Rate (ASR) by Temperature

