

Self-consistent errors are harder to detect

Qwen2.5-7B-Instruct / TriviaQA

