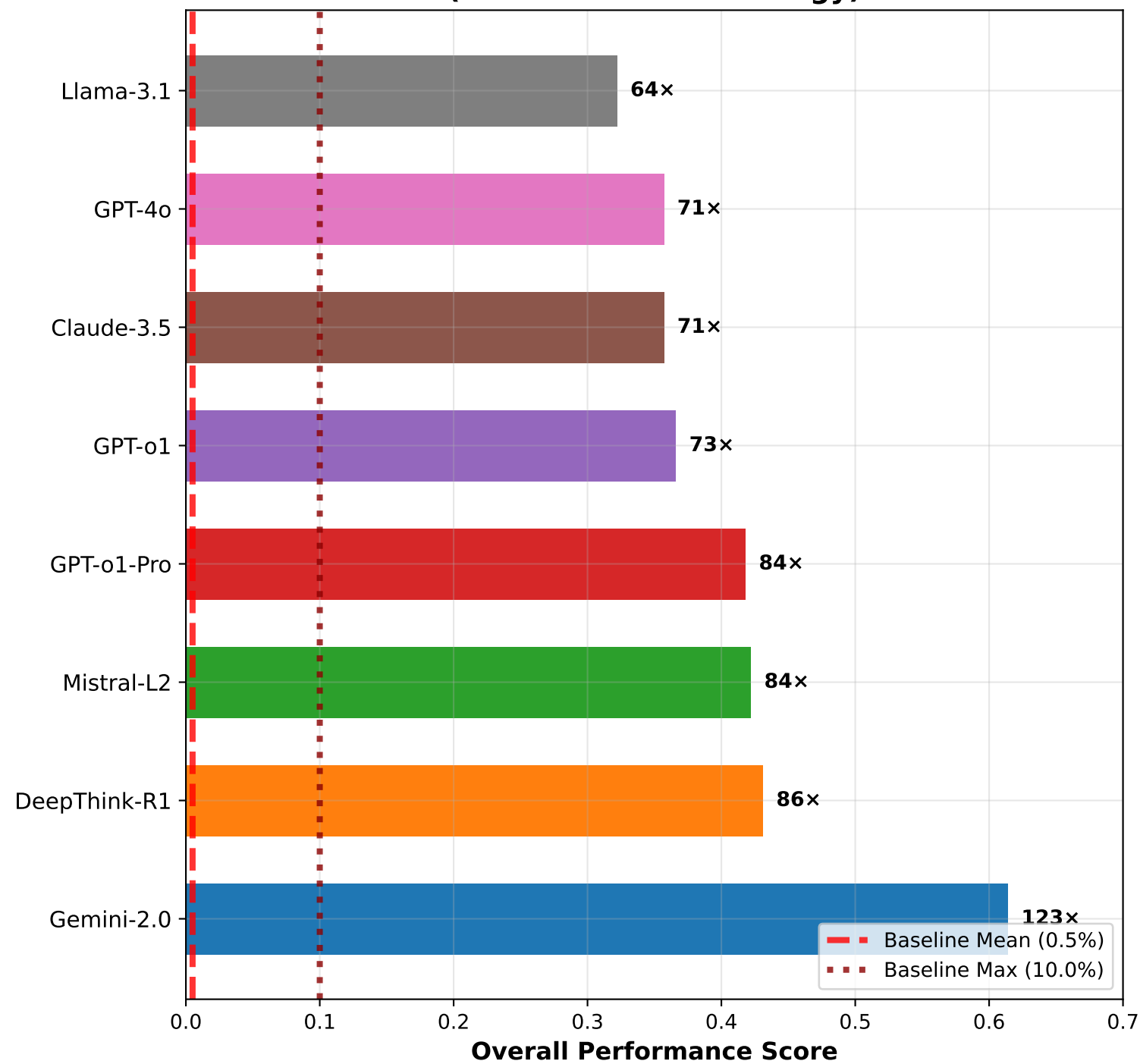


**LLM Performance vs Random Baseline
(Consistent Methodology)**



**Improvement Over Random Baseline
(64x to 123x Range)**

