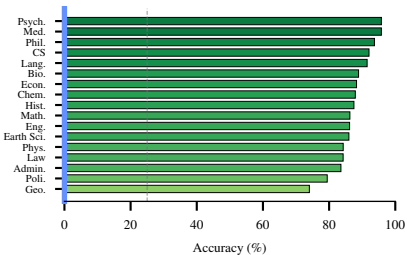
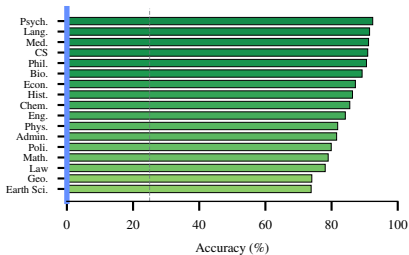


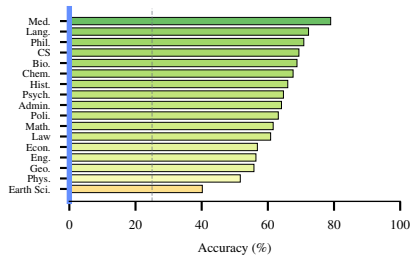
Gemini-2.5-pro (87.0%)



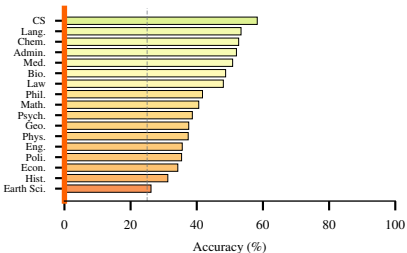
o3 (84.3%)



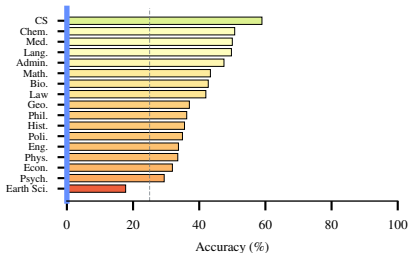
Claude-Sonnet-4 (63.3%)



Qwen2-VL-72B-Instruct (44.6%)



GPT-4o (42.0%)



InternVL2.5-38B-MPO (39.3%)

