

Model	Mode	Accuracy
vicuna-7b-v1.5	LLM	10.8
Llama-2-7b-chat-hf	LLM	34.6
MAP-Neo-7b-Instruct-v0.1	LLM	36.1
Meta-Llama3-8b-Instruct	LLM	43.7
Qwen2-7b-Instruct	LLM	44.7
Yi-1.5-6b-Chat	LLM	44.9
deepseek-llm-7b-chat	LLM	36.8
deepseek-vl-7b-chat	MLLM	60.3
deepseek-vl-7b-chat	text-only	36.8(- 23.5)
Qwen-VL-Chat	MLLM	53.4
Qwen-VL-Chat	text-only	38.0(- 15.4)
idefics2-8b	MLLM	67.7
idefics2-8b	text-only	47.8(- 19.9)
Mantis-8B-siglip-Llama3	MLLM	57.5
Mantis-8B-siglip-Llama3	text-only	43.3(- 14.2)
Yi-VL-6B-Chat	MLLM	61.3
Yi-VL-6B-Chat	text-only	44.9(- 16.4)
Yi-VL-34B-Chat	MLLM	67.9
Yi-VL-34B-Chat	text-only	52.2(- 15.7)