Experimental Results

Model	GSM8K %	HellaSwag %	MMLU %	MATH %	ARC-Challenge %	BBH-Boolean %
OpenO1-LLaMA-8B-v0.1	16	50	20	10	68	90
OpenO1-Qwen-7B-v0.1	20	68	46	2	88	96
SCoRe+OpenO1-Qwen-7B-v0.1	18	60	54	8	88	94
SCoRe+OpenO1-LLaMA-8B-v0.1	**21**	**65**	**48**	**12**	**90**	**95**