

Lasso pairwise classification accuracy, distinguishing each LLM from human text				
LLM	Trained on CAP		Trained on HAP-E	
	Training	Test on HAP-E	Training	Test on CAP
GPT-4o				
GPT-4o	95.9%	95.3%	96.5%	95.3%
GPT-4o Mini	96.7%	96.3%	96.9%	96.2%
Llama Instruct				
Llama 3 70B Instruct	91.1%	90.7%	92.1%	90.2%
Llama 3 8B Instruct	92.4%	91.8%	92.9%	91.9%
Llama Base				
Llama 3 70B	77.2%	73.2%	75.6%	75.3%
Llama 3 8B	77.6%	73.8%	76.3%	75.3%