

# RAID Benchmark Leaderboard

These leaderboards contain the test-set scores of various detector models. To submit your own model's predictions to the leaderboards, see [Leaderboard Evaluation](#).













**Domain** 
**Decoding Strategy** 
**Repetition Penalty** 
**Adversarial Attack** 
**Metric** 
39 entries match your current filters

Detector	Generator Model											
	Aggregate	chatgpt	gpt4	gpt3	gpt2	mistral	mistral-chat	cohere	cohere-chat	llama-chat	mpt	mpt-chat
Trinka AI	0.999	1.000	1.000	1.000	1.000	0.999	1.000	0.997	0.998	1.000	1.000	1.000
GiantMelon	0.999	1.000	1.000	1.000	1.000	0.999	1.000	0.997	0.998	1.000	0.999	1.000
Veredict Labs	0.999	1.000	1.000	1.000	0.999	0.999	1.000	0.997	0.998	1.000	0.999	1.000
Grammarly	0.999	0.999	0.999	1.000	0.999	0.999	0.999	0.997	0.998	1.000	0.999	0.999
SpeedAI	0.999	1.000	0.999	0.999	0.999	0.998	1.000	0.997	0.998	1.000	0.998	1.000
Zhuque AI Detection Assistant	0.999	0.999	0.999	0.999	0.999	0.998	0.999	0.994	0.998	0.999	0.999	0.999
QuillBot	0.998	1.000	0.999	0.999	0.999	0.998	1.000	0.990	0.994	1.000	0.999	1.000
TMR (Target Mining RoBERTa) - AI Text Detector	0.997	0.998	0.998	0.998	0.998	0.996	0.998	0.990	0.992	0.999	0.997	0.998
MELD	0.997	1.000	0.999	0.999	0.998	0.992	0.999	0.982	0.992	1.000	0.996	1.000
GPTZero	0.995	1.000	0.999	0.997	0.997	0.994	0.998	0.958	0.985	1.000	0.998	0.999
Tabularis AI Detector	0.994	0.998	0.995	0.995	0.995	0.991	0.999	0.977	0.990	0.999	0.991	0.998
Fast AI Detector	0.989	0.995	0.996	0.987	0.988	0.981	0.993	0.969	0.982	0.996	0.987	0.996
DeBERTa-ConPara	0.987	0.999	0.995	0.970	0.993	0.981	0.997	0.918	0.982	1.000	0.988	0.998
It's AI	0.985	0.999	0.992	0.984	0.990	0.979	0.998	0.921	0.977	0.999	0.976	0.996
GPTypo	0.983	0.994	0.988	0.986	0.986	0.980	0.991	0.932	0.974	0.992	0.978	0.993

Leaderboard - RAID Benchmark

Detector	Generator Model											
	Aggregate	chatgpt	gpt4	gpt3	gpt2	mistral	mistral-chat	cohere	cohere-chat	llama-chat	mpt	mpt-chat
ModernBERT AI Detection 	0.981	0.986	0.961	0.994	0.983	0.980	0.986	0.942	0.969	0.993	0.982	0.985
DeBERTa-ConPara-v2.2	0.974	0.994	0.990	0.917	0.986	0.975	0.988	0.864	0.957	0.998	0.979	0.992
WPF V141 Photon Fractal	0.972	0.980	0.986	0.977	0.966	0.955	0.972	0.920	0.963	0.980	0.979	0.993
DeBERTa-ConPara-v2.6	0.964	0.990	0.973	0.921	0.967	0.961	0.982	0.867	0.952	0.987	0.968	0.979
Kareem Elsamadicy (Independent Researcher)	0.962	0.979	0.983	0.957	0.942	0.935	0.979	0.934	0.963	0.995	0.934	0.984
e5-small-lora 	0.959	0.981	0.980	0.966	0.956	0.936	0.974	0.879	0.950	0.978	0.946	0.981
ADAL 	0.957	0.970	0.968	0.969	0.954	0.945	0.969	0.933	0.956	0.969	0.933	0.970
DivEye 	0.957	0.979	0.937	0.986	0.974	0.917	0.959	0.923	0.964	0.981	0.941	0.968
Desklib AI Text Detector v1.01 	0.952	0.982	0.962	0.913	0.976	0.932	0.978	0.815	0.917	0.981	0.953	0.981
BERT-tiny-4M 	0.952	0.971	0.976	0.933	0.959	0.947	0.958	0.836	0.921	0.972	0.955	0.983
Desklib 	0.950	0.984	0.948	0.917	0.966	0.930	0.985	0.792	0.926	0.990	0.939	0.978
AI Detector.review 	0.885	0.919	0.912	0.876	0.866	0.855	0.905	0.804	0.868	0.915	0.874	0.915
SuperAnnotate AI Detector 	0.878	0.988	0.985	0.914	0.757	0.688	0.979	0.828	0.910	0.988	0.772	0.970
RADAR 	0.791	0.931	0.918	0.916	0.694	0.637	0.866	0.668	0.837	0.856	0.656	0.881
warm_silence_129	0.785	0.885	0.890	0.727	0.760	0.693	0.866	0.694	0.813	0.869	0.621	0.856
RoBERTa-large (GPT2) 	0.756	0.734	0.696	0.788	0.925	0.743	0.745	0.680	0.758	0.692	0.792	0.698
RoBERTa-base (GPT2) 	0.744	0.722	0.646	0.778	0.873	0.724	0.739	0.671	0.754	0.729	0.772	0.703
GLTR 	0.728	0.854	0.731	0.781	0.796	0.613	0.821	0.651	0.787	0.878	0.529	0.651
RoBERTa (ChatGPT) 	0.687	0.775	0.723	0.779	0.694	0.570	0.778	0.733	0.756	0.784	0.426	0.708

Leaderboard - RAID Benchmark

Detector	Generator Model											
	Aggregate	chatgpt	gpt4	gpt3	gpt2	mistral	mistral-chat	cohere	cohere-chat	llama-chat	mpt	mpt-chat
Gaussian Extreme  	0.501	0.503	0.501	0.503	0.500	0.500	0.502	0.501	0.501	0.499	0.502	0.503
Luminar  	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
LLMDet  	0.499	0.380	0.381	0.345	0.616	0.523	0.482	0.421	0.467	0.688	0.611	0.325
Binoculars   	--	0.976	0.842	0.971	--	0.684	0.922	0.883	0.949	0.976	0.531	0.774
FastDetectGPT   	--	0.910	0.770	0.908	--	0.690	--	0.852	0.905	0.930	0.535	0.691