Model Performance - Balanced Accuracy

Phenotype

Model	Spore Form.	Motility	Biosafety	Animal Path.	Plant Path.	Host Assoc.	Extreme Env.	Cell Shape	Gram Stain	Biofilm
gemini-2.5-pro	0.87	0.88	0.87	0.79	0.72	0.77	0.72	0.71	0.69	0.55
deepseek-r1	0.94	0.88	0.91	0.78	0.74	0.72	0.73	0.72	0.65	0.58
claude-3.5-sonnet	0.94	0.90	0.90	0.80	0.76	0.73	0.72	0.71	0.65	0.51
grok-3-mini	0.95	0.88	0.90	0.82	0.78	0.71	0.72	0.71	0.65	0.51
gpt-5	0.95	0.91	0.87	0.78	0.71	0.79	0.71	0.72	0.65	0.51
gpt-4o	0.93	0.86	0.87	0.82	0.73	0.81	0.66	0.71	0.64	0.55
gemini-pro-1.5	0.97	0.76	0.89	0.82	0.79	0.76	0.63	0.71	0.64	0.57
gpt-oss-120b	0.96	0.86	0.77	0.74	0.71	0.80	0.71	0.71	0.65	0.59
o1-mini	0.94	0.86	0.84	0.77	0.73	0.76	0.69	0.71	0.67	0.52
claude-sonnet-4	0.95	0.87	0.85	0.82	0.75	0.68	0.70	0.71	0.65	0.50
gemini-2.0-flash-lite-001	0.96	0.79	0.82	0.79	0.75	0.77	0.67	0.71	0.65	0.54
gpt-4.1-nano	0.80	0.83	0.87	0.75	0.69	0.77	0.68	0.92	0.58	0.55
gpt-4	0.92	0.84	0.78	0.75	0.74	0.79	0.72	0.63	0.65	0.62
gpt-5-nano	0.91	0.86	0.84	0.72	0.73	0.78	0.69	0.71	0.68	0.50
deepseek-chat-v3-0324	0.91	0.87	0.85	0.76	0.75	0.76	0.69	0.69	0.65	0.53
gpt-oss-20b	0.91	0.87	0.83	0.73	0.72	0.78	0.66	0.67	0.66	0.55
llama-3-70b-instruct:nitro	0.92	0.80	0.83	0.83	0.75	0.66	0.63	0.71	0.65	0.50
gemini-flash-1.5	0.95	0.82	0.82	0.85	0.78	0.63	0.63	0.62	0.63	0.50
wizardlm-2-8x22b	0.89	0.84	0.76	0.80	0.75	0.69	0.69	0.70	0.64	0.50
claude-3-haiku:beta	0.92	0.82	0.82	0.84	0.74	0.78	0.53	0.66	0.64	0.50
gemini-pro	0.89	0.81	0.88	0.83	0.67	0.73	0.60	0.65	0.62	0.50
gemini-2.5-flash-lite	0.91	0.86	0.74	0.71	0.75	0.54	0.68	0.70	0.64	0.54
gpt-3.5-turbo-0125	0.91	0.79	0.66	0.74	0.73	0.70	0.63	0.68	0.64	0.50
palm-2-chat-bison-32k	0.86	0.76	0.77	0.79	0.70	0.76	0.61	0.58	0.62	0.53
mixtral-8x7b-instruct:nitro	0.92	0.78	0.58	0.70	0.71	0.59	0.62	0.57	0.61	0.50
mistral-7b-instruct	0.86	0.69	0.71	0.74	0.69	0.65	0.57	0.49	0.49	0.61
openchat-7b	0.76	0.70	0.52	0.74	0.68	0.69	0.57	0.50	0.68	0.59
llama-3-8b-instruct:nitro	0.86	0.77	0.54	0.78	0.69	0.60	0.57	0.47	0.54	0.50
llama-3-sonar-small-32k-chat	0.88	0.64	0.61	0.76	0.66	0.50	0.63	0.40	0.58	0.50
phi-3-mini-128k-instruct	0.72	0.70	0.71	0.59	0.62	0.58	0.57	0.50	0.51	0.48
gemma-7b-it	0.67	0.52	0.54	0.69	0.66	0.58	0.50	0.28	0.43	0.50