Confusion Matrices for 'llama4'

gemini vs Human

	Model Score 4	Model Score 5
Human Score 4	0.90	0.10
Human Score 5	0.68	0.32

prometheus vs Human

	Model Score 4	Model Score 5
Human Score 4	0.00	1.00
Human Score 5	0.00	1.00