Confusion (One-vs-Rest): SUPPORTED (normalize=true)

True (Human) \ Predicted (LLM)	SUPPORTED	NOT SUPPORTED
SUPPORTED	8 — 88.9%	1 — 11.1%
NOT SUPPORTED	10 — 30.3%	23 — 69.7%