

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

True (Human)

SUPPORTED

8

44.4%

10

55.6%

NOT SUPPORTED

1

4.2%

23

95.8%

SUPPORTED
NOT SUPPORTED

Predicted (LLM)