

Predicted Sample Type

gpt-4.1-nano

gpt-4.1-mini

gpt-4.1

o4-mini

o3

gemma-3-1b

1.0

0.5

0.0

positive

negative

unknown

0

25

50

0

25

50

0

25

50

0

25

50

0

25

50

0

25

50

Sample Length

