

Training Loss

2.25
2.00
1.75
1.50
1.25

0

100

200

300

400

500

Epoch



Llama-3.1 (8B)



gpt-oss (20B)



Llama-3.1 (70B)



gpt-oss (120B)

