

validation-set inference-aware loss

0.55  
0.50  
0.45  
0.40  
0.35  
0.30  
0.25  
0.20

0

5000

10000

15000

20000

training step

