

training cross entropy

2.0

1.5

1.0

0.5

0.0

0.0

1.6

3.1

4.7

6.3

7.8

9.4

10.9

12.5

14.1

15.6

steps (x 1000)

