recurrent network augmented transformer @ MemoryBenchmark 1.0 2.0 0.8 1.5 0.6 train sparse categorical crossentropy train sparse categorical accuracy val sparse categorical crossentropy val sparse categorical accuracy test sparse categorical crossentropy test sparse categorical accuracy 1.0 0.4 0.5 - 0.2 0.0

16

epochs

20

24

12

28