recurrent network attention transformer @ AddBenchmark 0.175 0.1500.125 0.100 train mean squared error val mean squared error test mean squared error 0.075 0.050 0.025 0.000 6 16 8 10 12 14 18

epochs