

000
001
002
003
004
005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

Feed-Forward Networks with Attention Can Solve Some Long-Term Memory Problems

Anonymous Author(s)

Affiliation

Address

email

Abstract

Recently, recurrent neural networks (RNNs) have been augmented with “attention” mechanisms which compute a fixed-length representation of entire sequences. We propose a simplified model of attention which is applicable to feed-forward neural networks and demonstrate that it can solve some long-term memory problems (specifically, those where temporal order doesn’t matter). In fact, we show empirically that our model can solve these problems for sequence lengths which are both longer and more widely varying than has been shown for RNNs.

1 Attention