

Reading Report of "Encoding Recurrent into Transformers"

from ICLR 2023

PB19030800 陈磊

Prerequisite knowledge

RNN

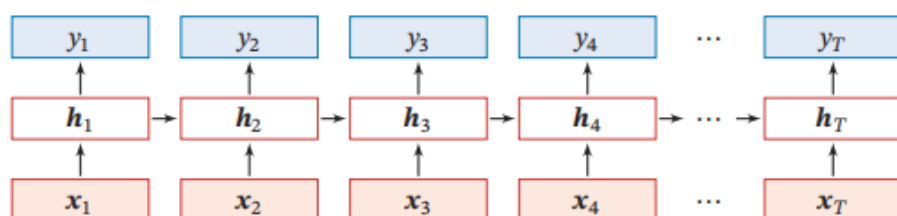


图 6.2 按时间展开的循环神经网络

$$h_t = f(Uh_{t-1} + Wx_t + b)$$
$$y_t = Vh_t$$

LSTM

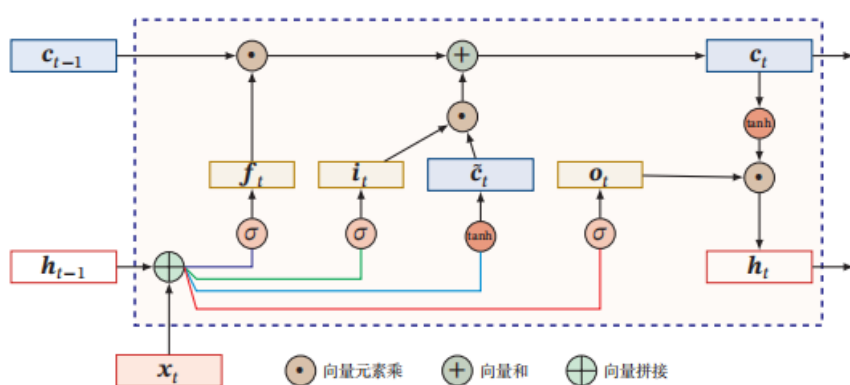


图 6.7 LSTM 网络的循环单元结构