

Recurrent Neural Network

Congshan Zhang

1 Unfolding a Recurrent Neural Network

The training set X contains N individuals and D features over T periods. So we can represent it as a three-dimensional tensor with shape $(N \times T \times D)$.

We have an unfolded RNN below in Figure 1.

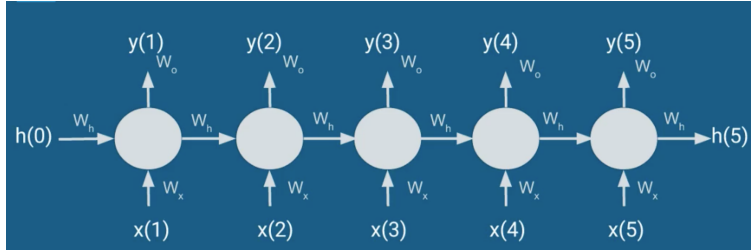


Figure 1: This is an illustration of a recurrent neural network with shared weighting.

1.1 Backpropagation Through Time (BPTT)

At time t , the output of the network is

$$\hat{y}(t) = \text{softmax}(W_o^T h(t)). \quad (1)$$

Applying the definition of $h(t)$ yields

$$\hat{y}(t) = \text{softmax}(W_o^T f(W_h^T h(t-1) + W_x^T x(t))). \quad (2)$$

One can write it recursively as

$$\begin{aligned} \hat{y}(t) &= \text{softmax}(W_o^T f(W_h^T f(W_h^T h(t-2) + W_x^T x(t-1)) + W_x^T x(t))) \\ &= \dots \end{aligned} \quad (3)$$

References