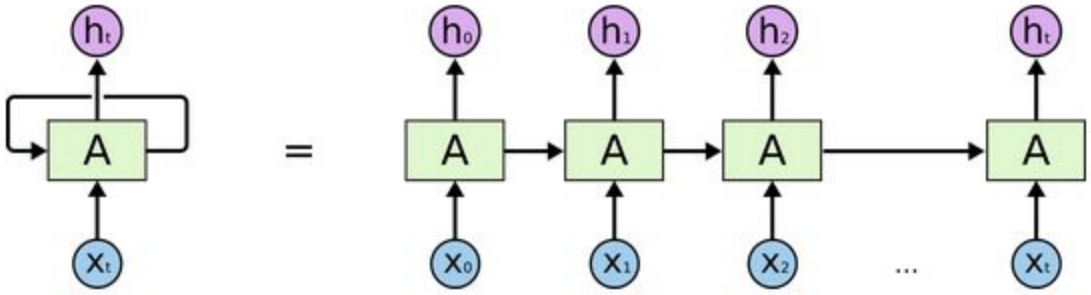
How Do Neural Nets Remember?

Recurrent Neural Networks:



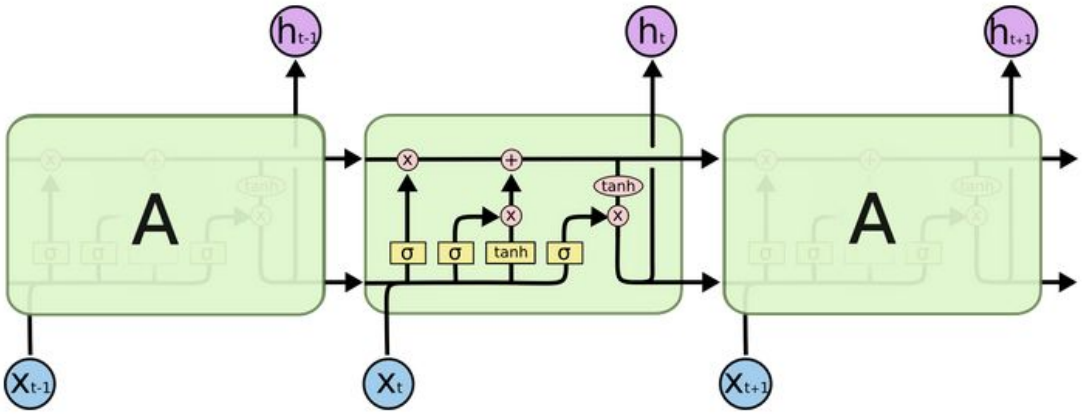
An unrolled recurrent neural network.

Recurrent Networks loop some information backward in order to help make predictions. For example, trying to predict the word *sky* is easy when the words before it are "The clouds are in the..". These networks are going to allow ordinal data to be easily interpreted. This powerful feature does however have its limitations.

The network can't effectively pass information down an extended chain. In other words, the network has short term memory but fails to remember long term trends.

Introducing…

Long Short Term Memory:



The repeating module in an LSTM contains four interacting layers.

Long Short Term Memory includes a much more advanced activation sequence:

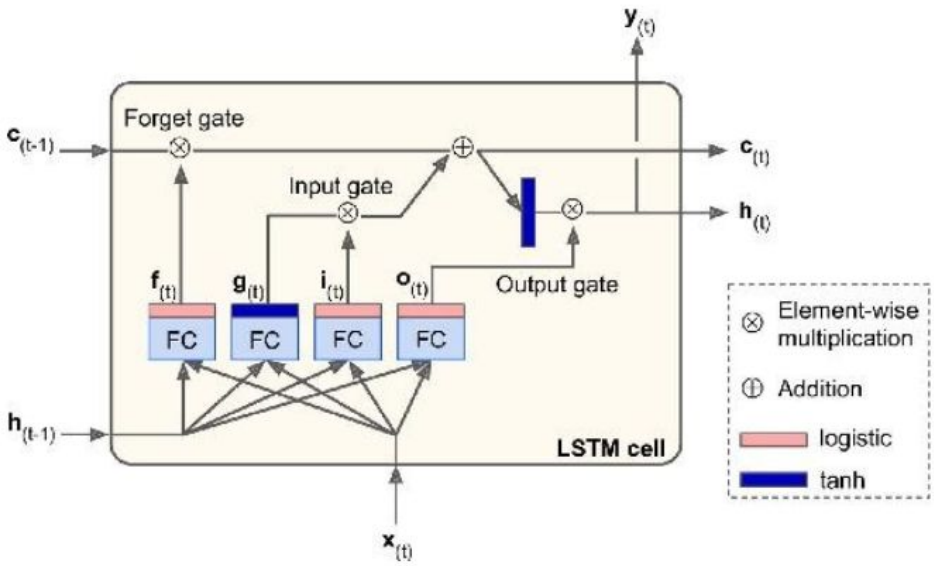


Figure 14-13. LSTM cell

Short Term State: **h**(t)
Long Term State: **c**(t)

- **Forget Gate** (controlled by f(t)) controls which parts of the long-term state should be erased.
- **Input Gate** (controlled by i(t)) controls which parts of g(t) should be added to the long-term state

- **Output Gate** (controlled by o(t)) controls which parts of the long-term state should be read and output at this time step (both to h(t)) and y(t)

What is being calculated:

$$\mathbf{i}_{(t)} = \sigma(\mathbf{W}_{xi}^{T} \cdot \mathbf{x}_{(t)} + \mathbf{W}_{hi}^{T} \cdot \mathbf{h}_{(t-1)} + \mathbf{b}_{i})$$

$$\mathbf{f}_{(t)} = \sigma(\mathbf{W}_{xf}^{T} \cdot \mathbf{x}_{(t)} + \mathbf{W}_{hf}^{T} \cdot \mathbf{h}_{(t-1)} + \mathbf{b}_{f})$$

$$\mathbf{o}_{(t)} = \sigma(\mathbf{W}_{xo}^{T} \cdot \mathbf{x}_{(t)} + \mathbf{W}_{ho}^{T} \cdot \mathbf{h}_{(t-1)} + \mathbf{b}_{o})$$

$$\mathbf{g}_{(t)} = \tanh(\mathbf{W}_{xg}^{T} \cdot \mathbf{x}_{(t)} + \mathbf{W}_{hg}^{T} \cdot \mathbf{h}_{(t-1)} +$$

$$\mathbf{c}_{(t)} = \mathbf{f}_{(t)} \otimes \mathbf{c}_{(t-1)} + \mathbf{i}_{(t)} \otimes \mathbf{g}_{(t)}$$

$$\mathbf{y}_{(t)} = \mathbf{h}_{(t)} = \mathbf{o}_{(t)} \otimes \tanh(\mathbf{c}_{(t)})$$