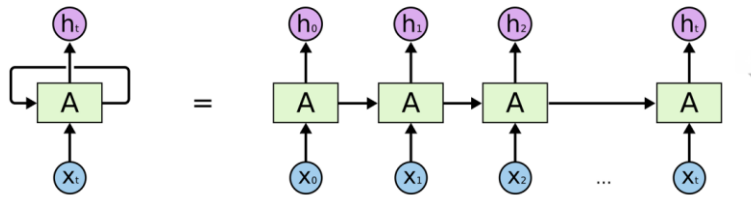


Understanding LSTM Networks

Tiancheng Yuan & Renzhi Sheng

The strengths of Recurrent Neural Networks(RNN)

- The networks with loops in them, allow information to persist
- Be thought of as multiple copies of the same network, each passing a message to a successor
- Connect previous information to the present task



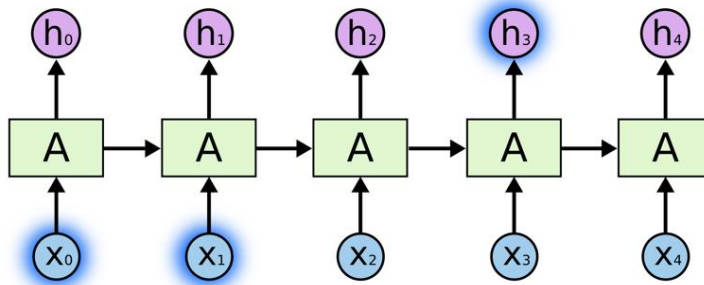
An unrolled recurrent neural network.

In the above diagram, a chunk of neural network, A , looks at some input x_t and outputs a value h_t . A loop allows information to be passed from one step of the network to the next.

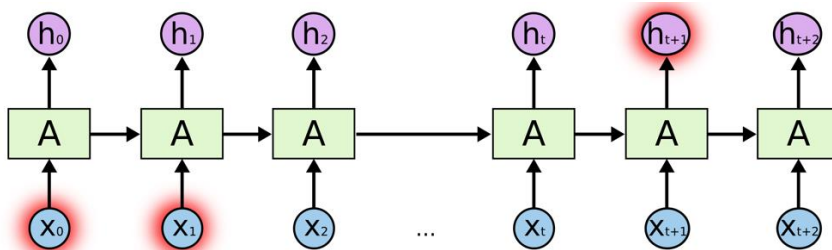
This chain-like nature reveals that recurrent neural networks are intimately related to sequences and lists. They're the natural architecture of neural network to use for such data.

The weaknesses of Recurrent Neural Networks(RNN)

- Be unable to learn to connect the information as the gap between the relevant information grows



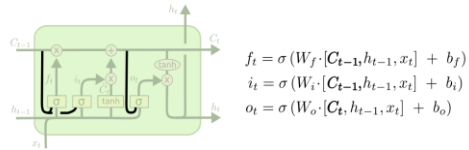
- where the gap between the relevant information and the place that it's needed is small, RNNs can learn to use the past information.



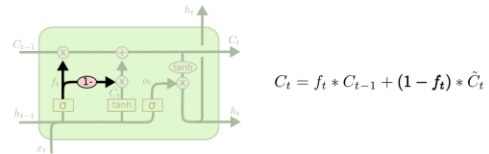
- But there are also cases where we need more context. Consider trying to predict the last word in the text “I grew up in France... I speak fluent French.” Recent information suggests that the next word is probably the name of a language, but if we want to narrow down which language, we need the context of France, from further back. It's entirely possible for the gap between the relevant information and the point where it is needed to become very large.
- Unfortunately, as that gap grows, RNNs become unable to learn to connect the information.
- However, LSTMs don't have this problem.

Variants on LSTM

- Adding “ peephole connections”



- Use coupled forget and input gates



- The Gated Recurrent Unit, or GRU

