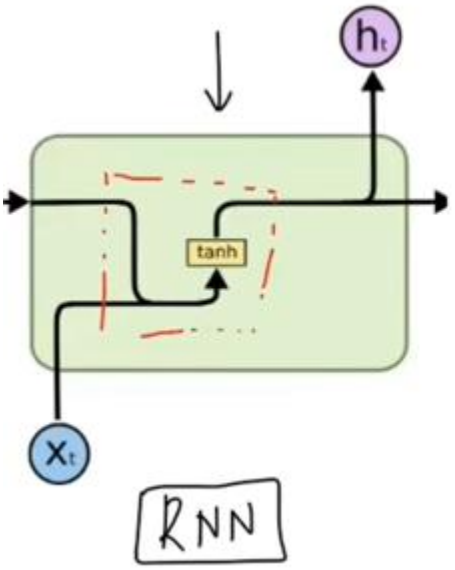
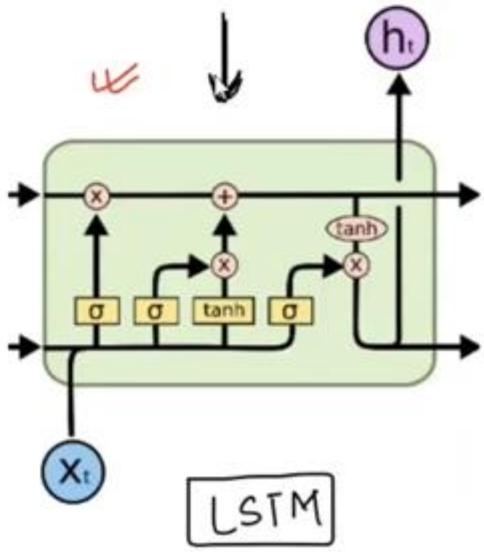
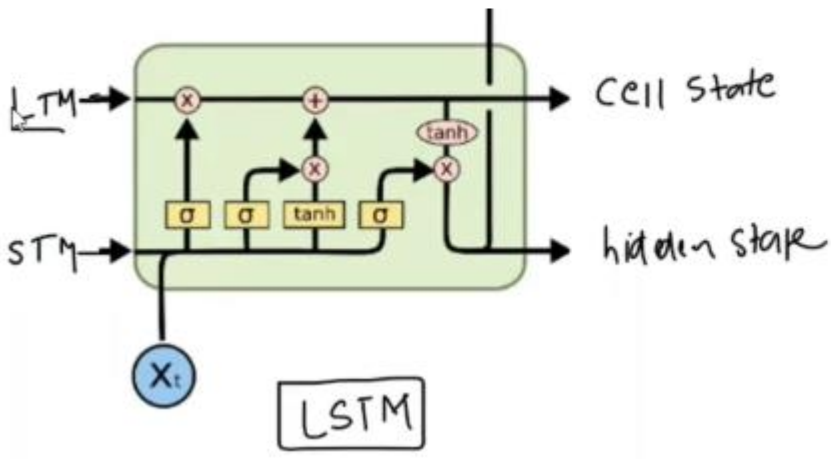
How of LSTM

Architecture







Long Term Memory & ShortTerm Memory Dono State maintain krrha

Gates of LSTM

[Understanding LSTM Networks -- colah's blog](https://colah.github.io/posts/2015-08-Understanding-LSTMs/)

https://colah.github.io/posts/2015-08-Understanding-LSTMs/

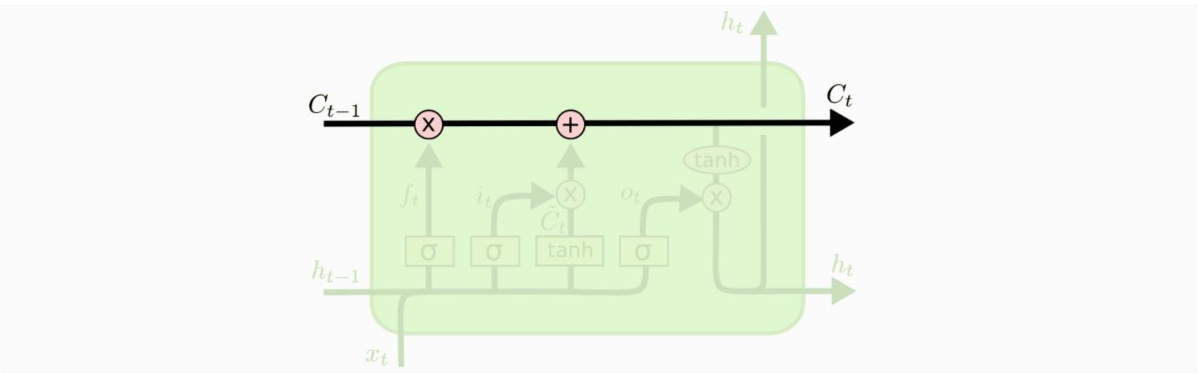
The Core idea Behind LSTM

The key to LSTMs is the cell state, the horizontal line running through the top of Diagram,

The cell state is kind of like a conveyor belt.

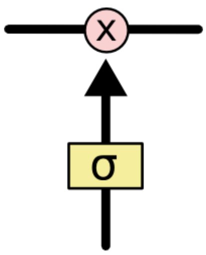
It runs straight down the entire chain, with only some minor linear interactions.

It's very easy for information to just flow along it unchanged.



The LSTM does have the ability to remove or add information to the cell state, carefully regulated by structures called gates.

Gates are a way to optionally let information through, They are composed out of a sigmoid neural net layer and a pointwise multiplication operation.



The sigmoid layer outputs number between 0 and 1, describing how much of each component should be let through.

A Value of zero means “let nothing Through or block almost everything”.

An LSTM has 3 of these Gates, to protect and control the cell state.

Conclusion

Earlier, I mentioned the remarkable results people are achieving with RNNs. Essentially all of these are achieved using LSTMs. They really work a lot better for most tasks!

Written down as a set of equations, LSTMs look pretty intimidating. Hopefully, walking through them step by step in this essay has made them a bit more approachable.

LSTMs were a big step in what we can accomplish with RNNs. It’s natural to wonder: is there another big step?

A common opinion among researchers is: “Yes! There is a next step and it’s attention!” The idea is to let every step of an RNN pick information to look at from some larger collection of information.

For example, if you are using an RNN to create a caption describing an image, it might pick a part of the image to look at for every word it outputs. In fact, Xu, *et al.* (2015) do exactly this – it might be a fun starting point if you want to explore attention! There’s been a number of really exciting results using attention, and it seems like a lot more are around the corner…