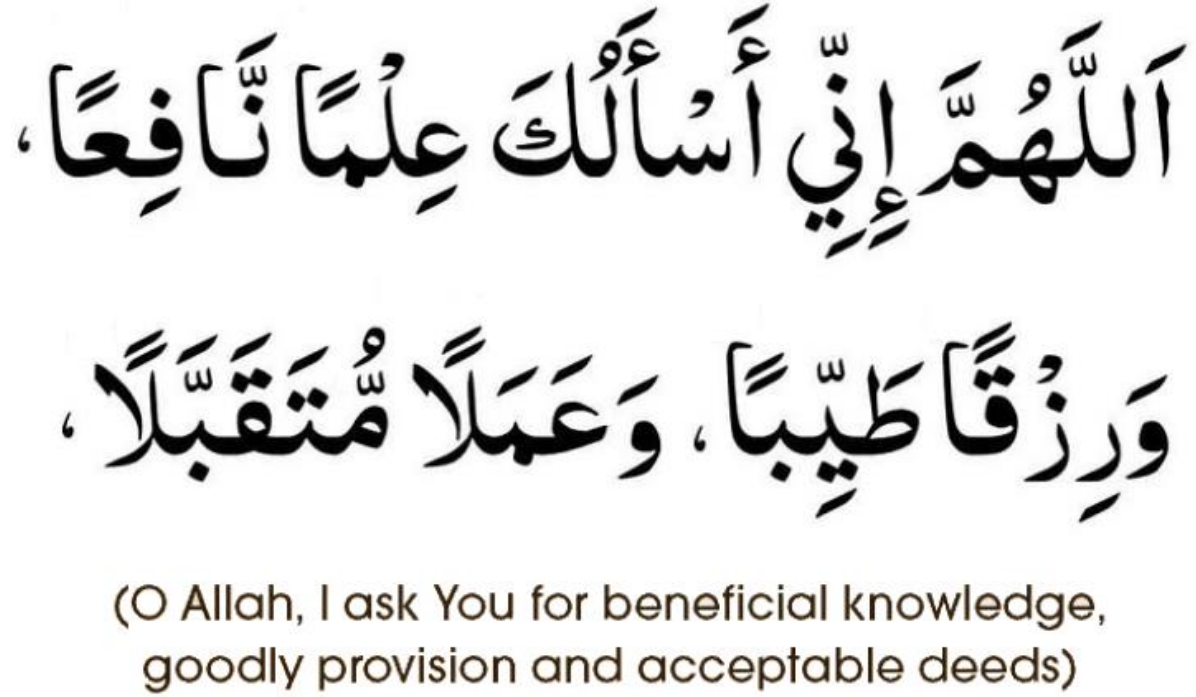
Artificial Intelligence

Unit 10
LSTM and GRU

By:
Syeda Saleha Raza

AL NAFI,
A company with a focus on education,
wellbeing and renewable energy.



اے اللہ ، میں آپ سے سوال کرتی ہوں نفع بخش علم کا، طیّب رزق کا، اور اس عمل کا

(Sunan Ibn Majah: 925)

Outline

- Intuition
- RNN Recap
- Introducing LSTM
- Gates in LSTM
- Gated Recurrent Unit (GRU)
- Hands-on

LSTM - Intuition

Intuition

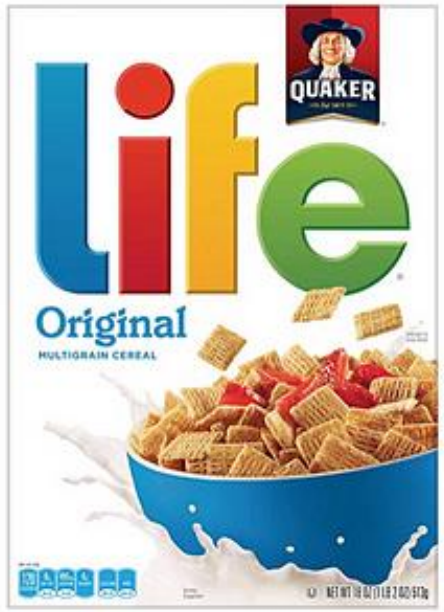
Customers Review 2,491

Thanos

September 2018

Verified Purchase

Amazing! This box of cereal gave me a perfectly balanced breakfast, as all things should be. I only ate half of it but will definitely be buying again!



A Box of Cereal

$3.99

Customers Review 2,491

Thanos

September 2018

Verified Purchase

Amazing! This box of cereal gave me a perfectly balanced breakfast, as all things should be. I only ate half of it but will definitely be buying again!

A Box of Cereal

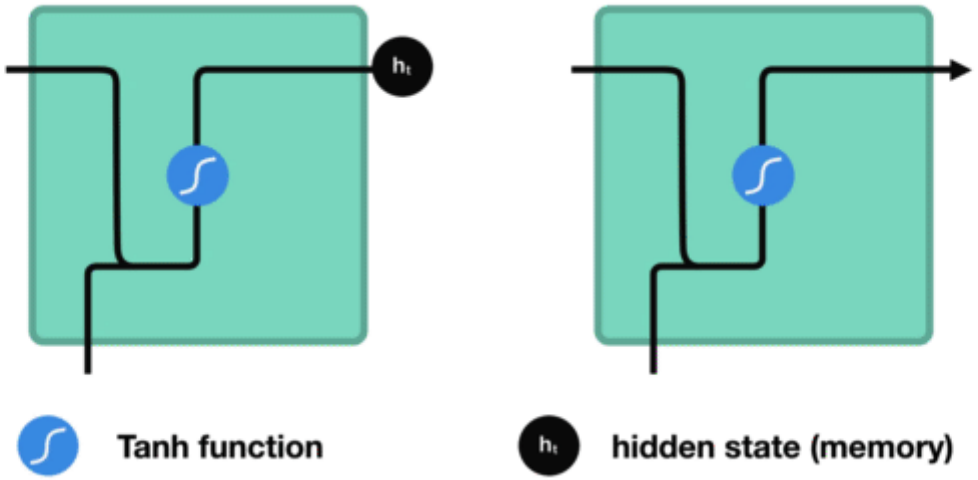
$3.99

- This is essentially what an LSTM or GRU does. It can learn to keep only relevant information to make predictions, and forget non relevant data.

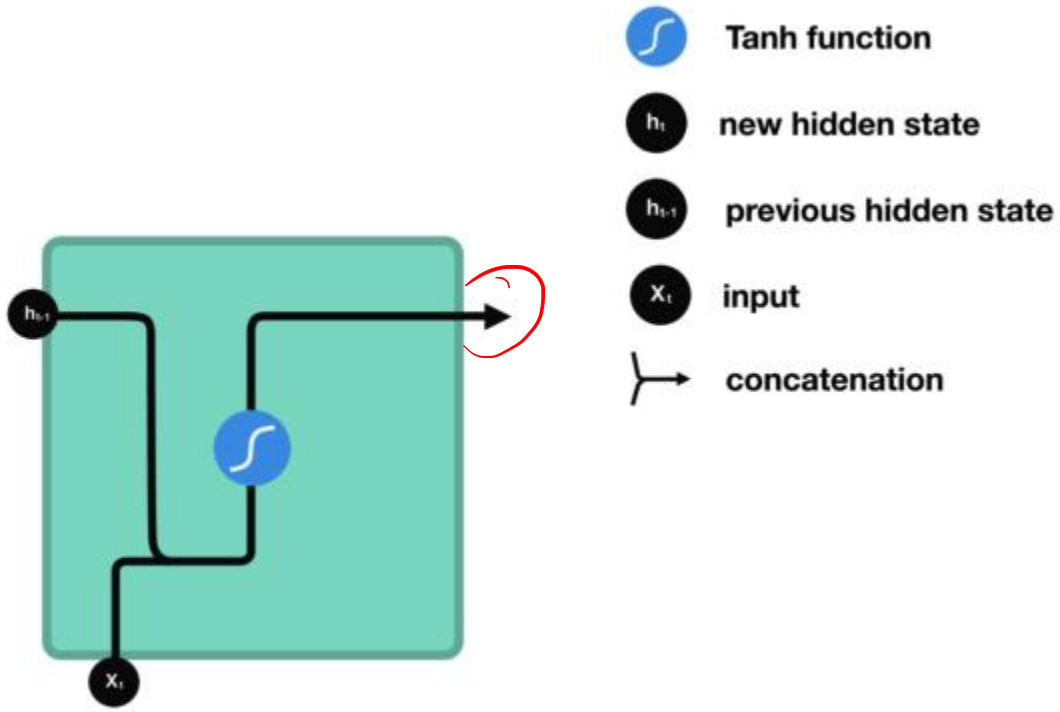
RNN - Recap

Review of RNN

Passing hidden state to next time step



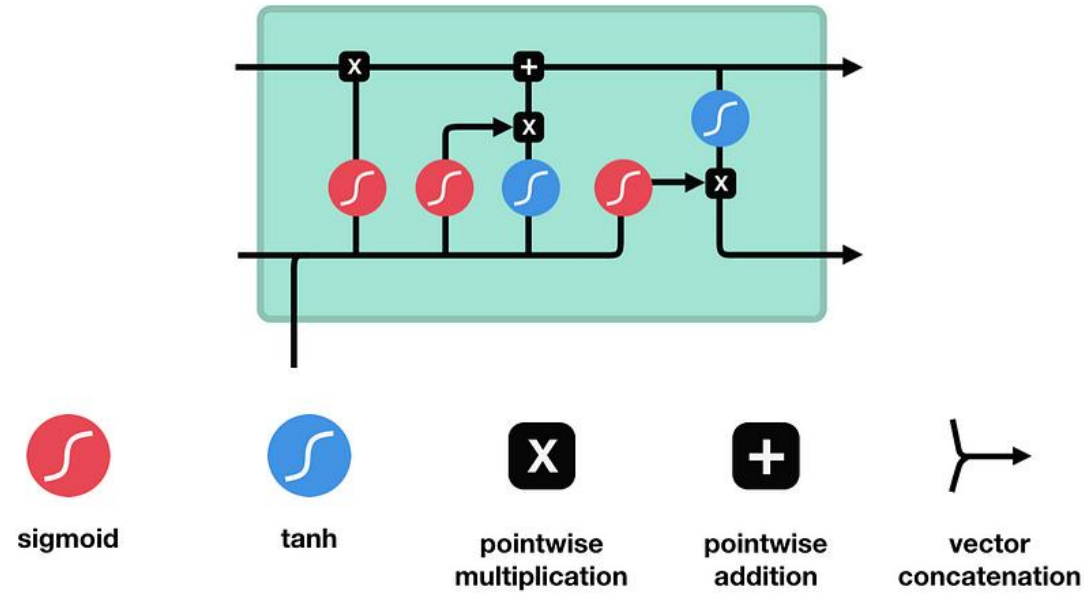
RNN Cell



LSTM Cell

LSTM Cell

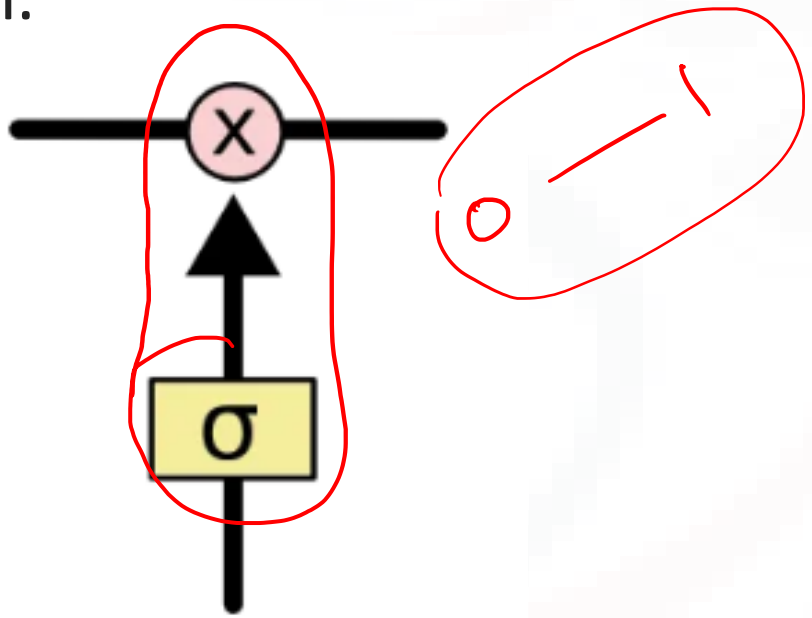
- An LSTM has a similar control flow as a recurrent neural network. It processes data passing on information as it propagates forward. The differences are the operations within the LSTM's cells.



These operations are used to allow the LSTM to keep or forget information.

Gates

- **Gates** are a way to optionally let information through. They are composed out of a sigmoid neural net layer and a point-wise multiplication operation.



LSTM Gate

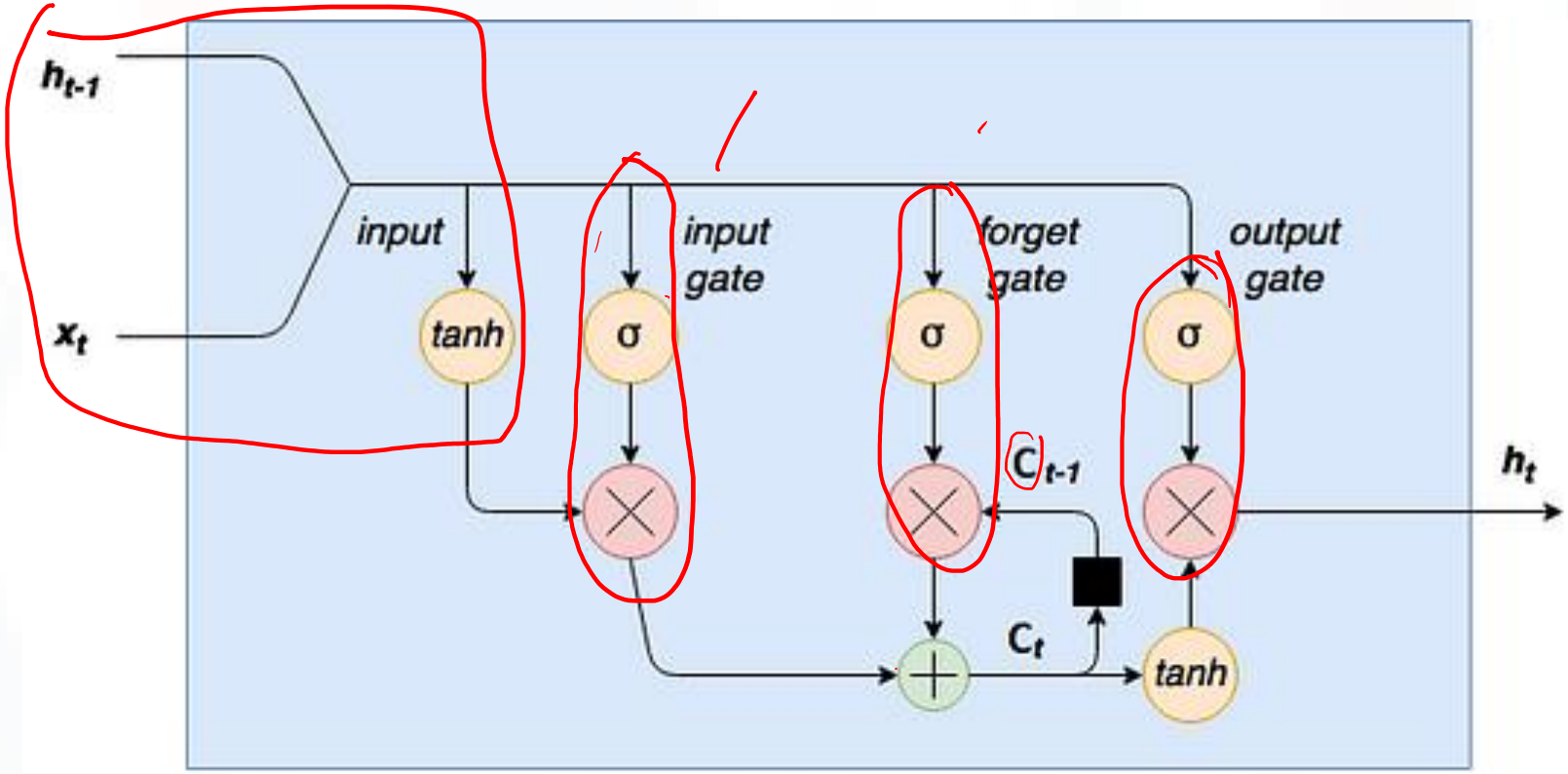
Cell State

- Memory of the network
- transfers relative information all the way down the sequence chain
- even information from the earlier time steps can make it’s way to later time steps.
- As the cell state goes on its journey, information get’s added or removed to the cell state via gates.

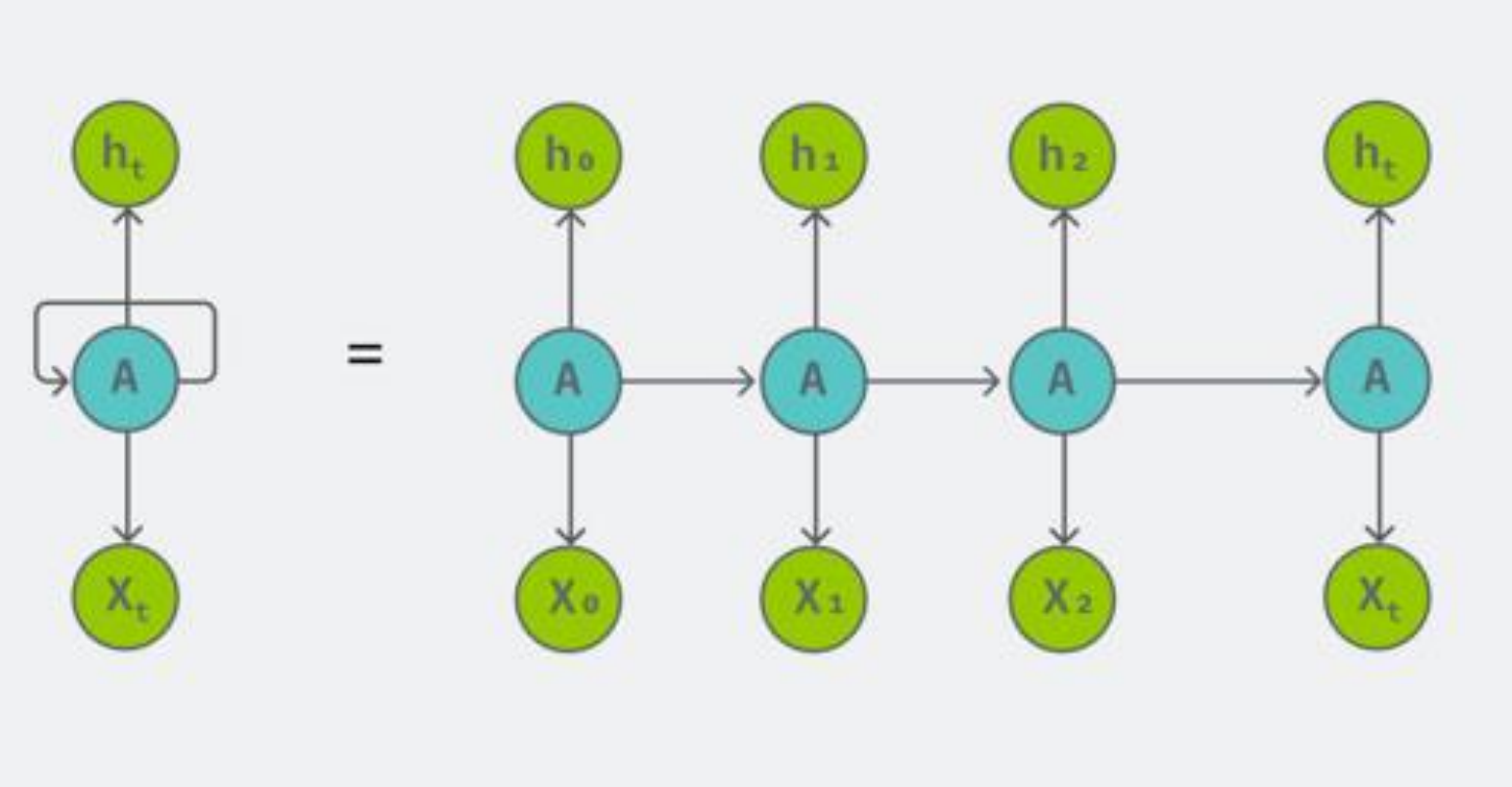
Gates

- An LSTM neuron can do this learning by incorporating a *cell state* and three different *gates*:
  - *input gate*,
  - *forget gate* and
  - *output gate.*
- The input gate decides whether to update the cell state or not.
- The forget gate can erase cell memory
- The output gate can decide whether to make the output information available or not.

Summarizing LSTM cell



LSTM Architecture



Reducing vanishing gradient

- The LSTM cell is a specifically designed unit of logic that will help reduce the vanishing gradient.
- The way it does so is by creating an internal memory state which is simply *added* to the processed input, which greatly *reduces the multiplicative effect of small gradients*.
- The time dependence and effects of previous inputs are controlled by an interesting concept called a *forget gate*, which determines which states are remembered or forgotten.

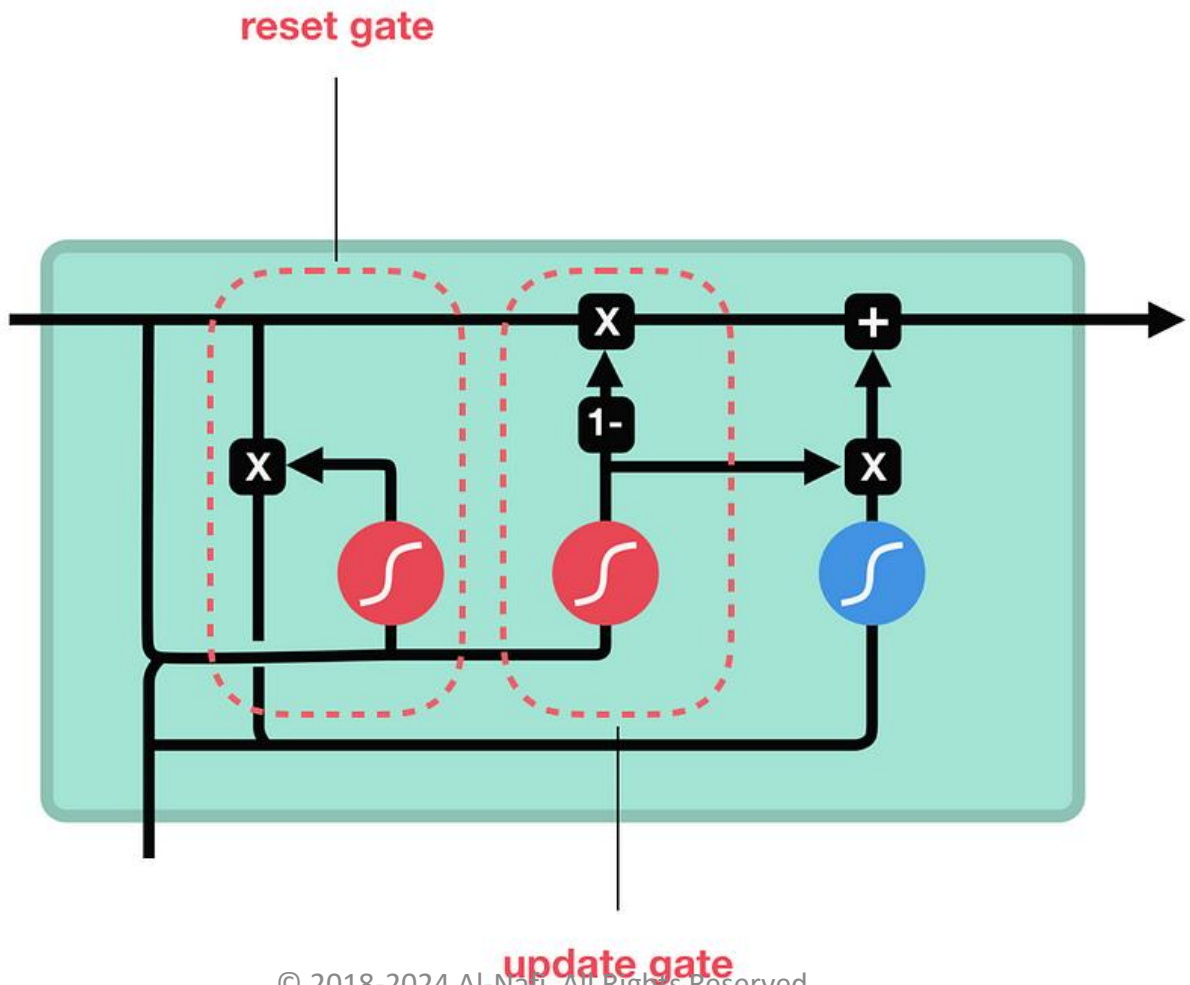
Code Demo – LSTM and GRU

- Adding LSTM and GRU layers in our Keras model
-Language generation using LSTM/GRU

Gated Recurrent Unit (GRU)

- The GRU is pretty similar to an LSTM. GRU's got rid of the cell state and used the hidden state to transfer information. It also only has two gates, a reset gate and update gate.
- **Update Gate**
  - The update gate acts similar to the forget and input gate of an LSTM. It decides what information to throw away and what new information to add.
- **Reset Gate**
  - The reset gate is another gate is used to decide how much past information to forget.

Gated Recurrent Unit (GRU)



Code Demo - GRU

References

- https://www.v7labs.com/blog/recurrent-neural-networks-guide
- https://medium.com/@humble_bee/rnn-recurrent-neural-networks-lstm-842ba7205bbf
- Illustrated Guide to Recurrent Neural Networks | by Michael Phi | Towards Data Science

جزاك اللهُ

To ask questions, Please use communities link (for respective course) within portal

https://portal.alnafi.com/enrollments