



Deep Learning

Prof. Silvio R. R. Sanches

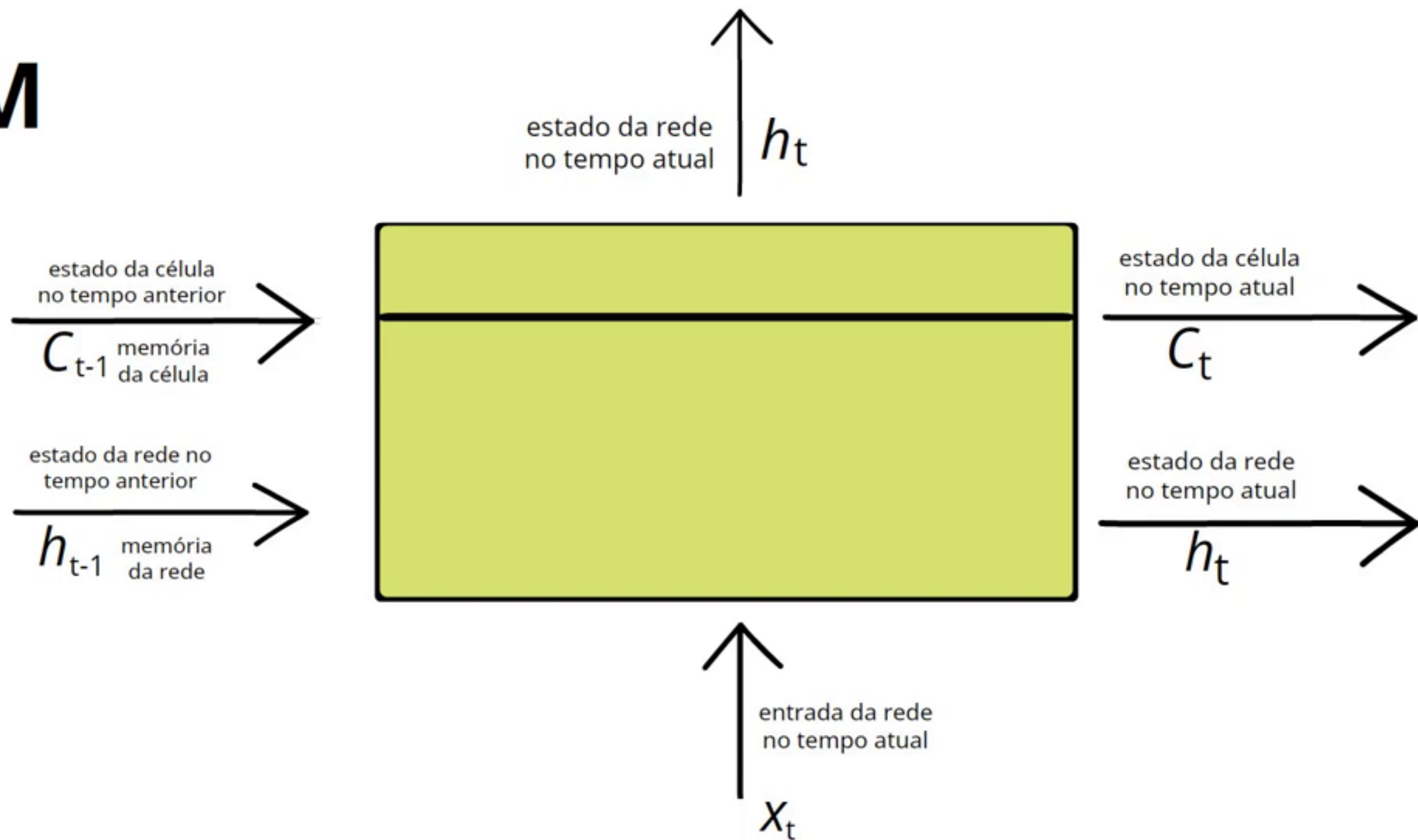
silviosanches@utfpr.edu.br

UTFPR

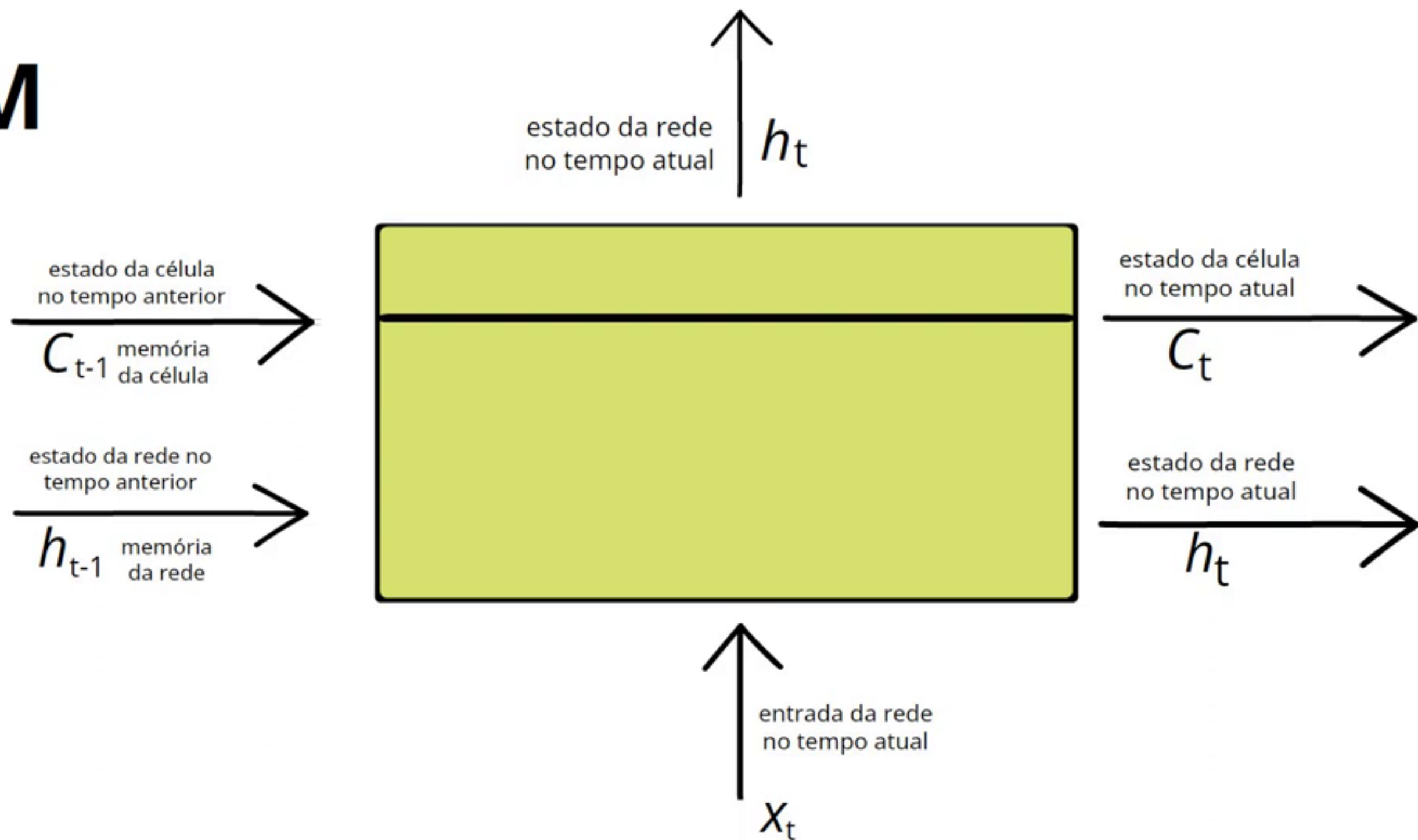
Redes Neurais Recorrentes

Memória Longa de Curto Prazo

LSTM

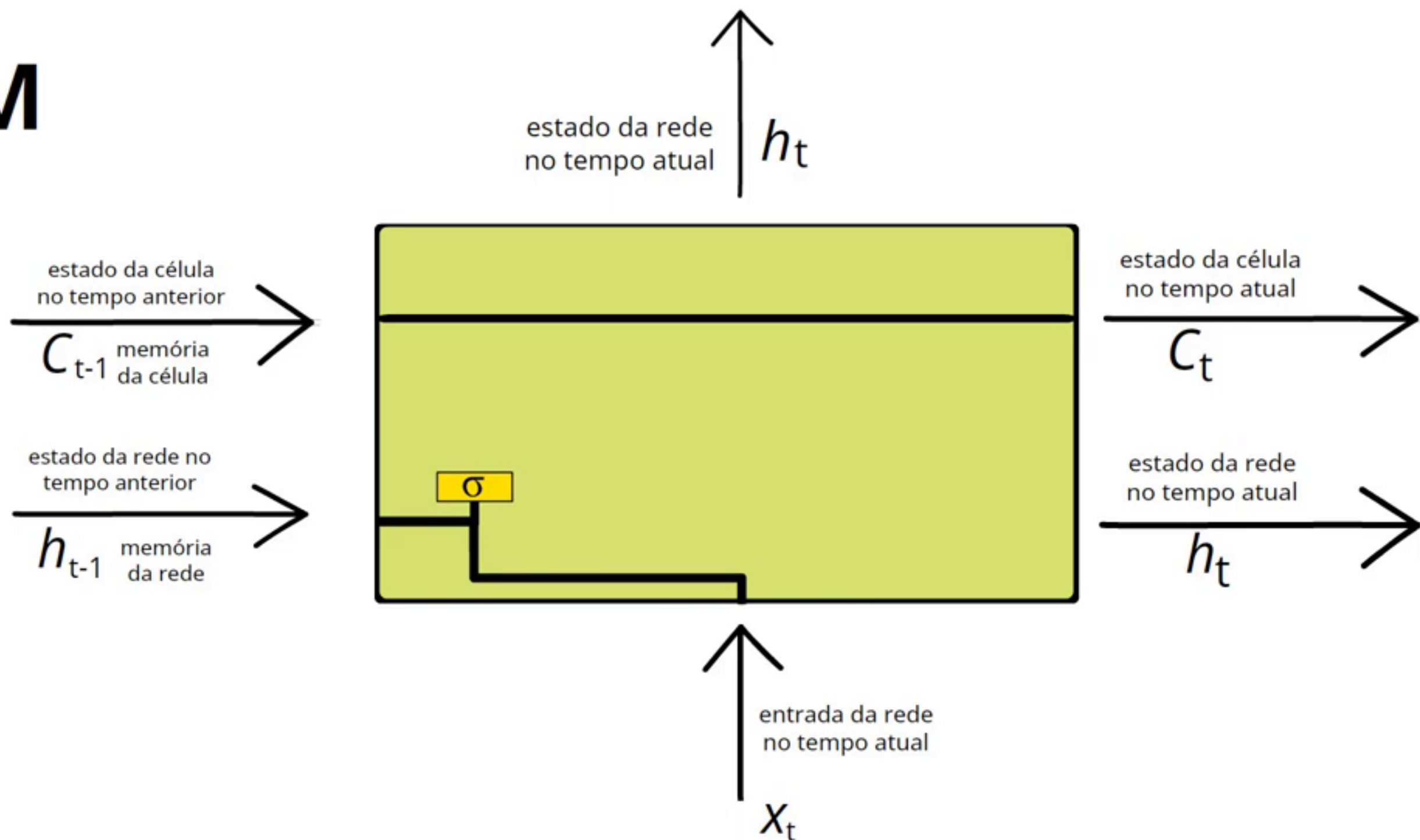


LSTM



Inicialmente, informações do estado oculto no tempo anterior e a entrada no tempo atual passam por uma camada de rede linear e o resultado dessa combinação passa por uma função sigmóide

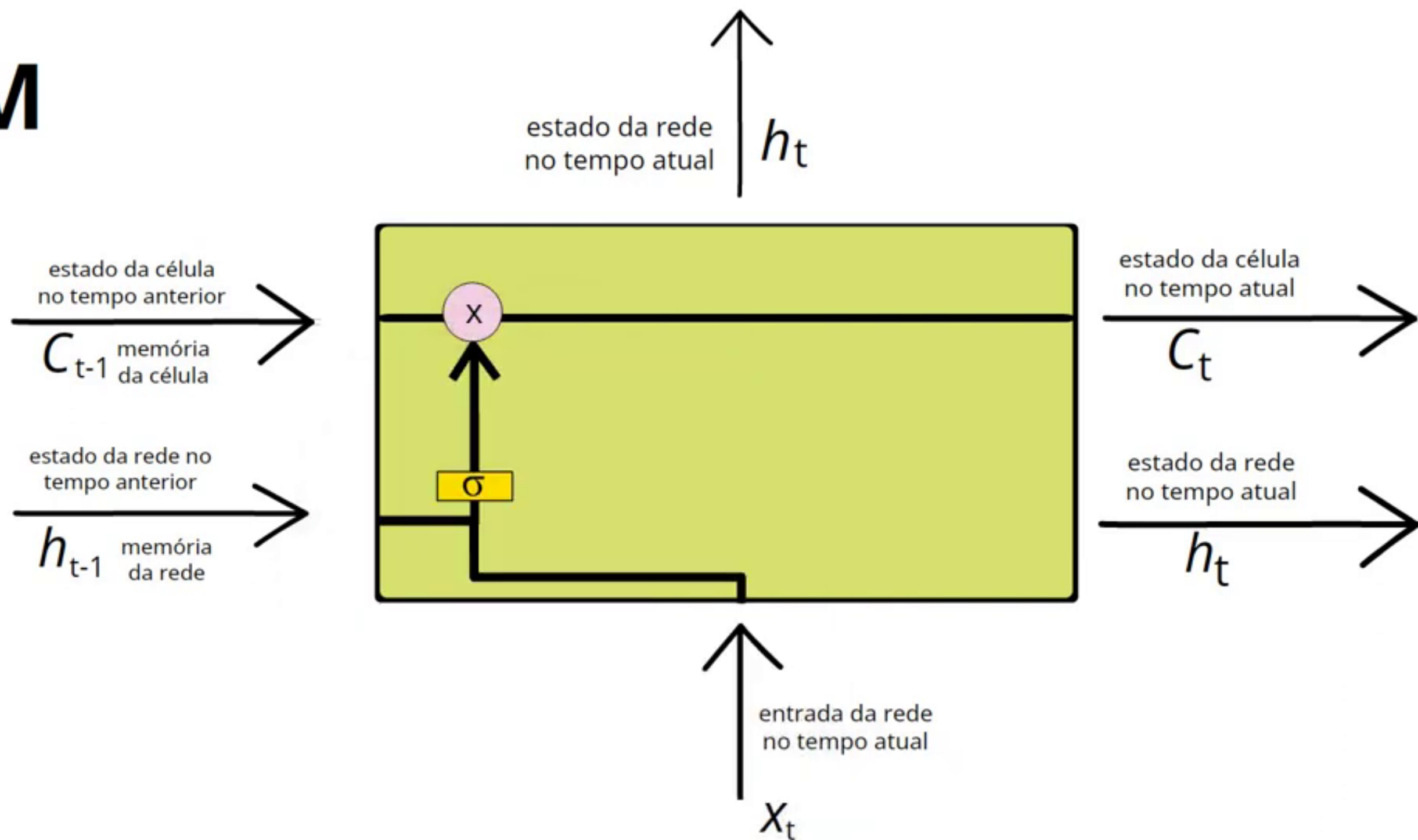
LSTM



σ Retorna valores entre 0 e 1

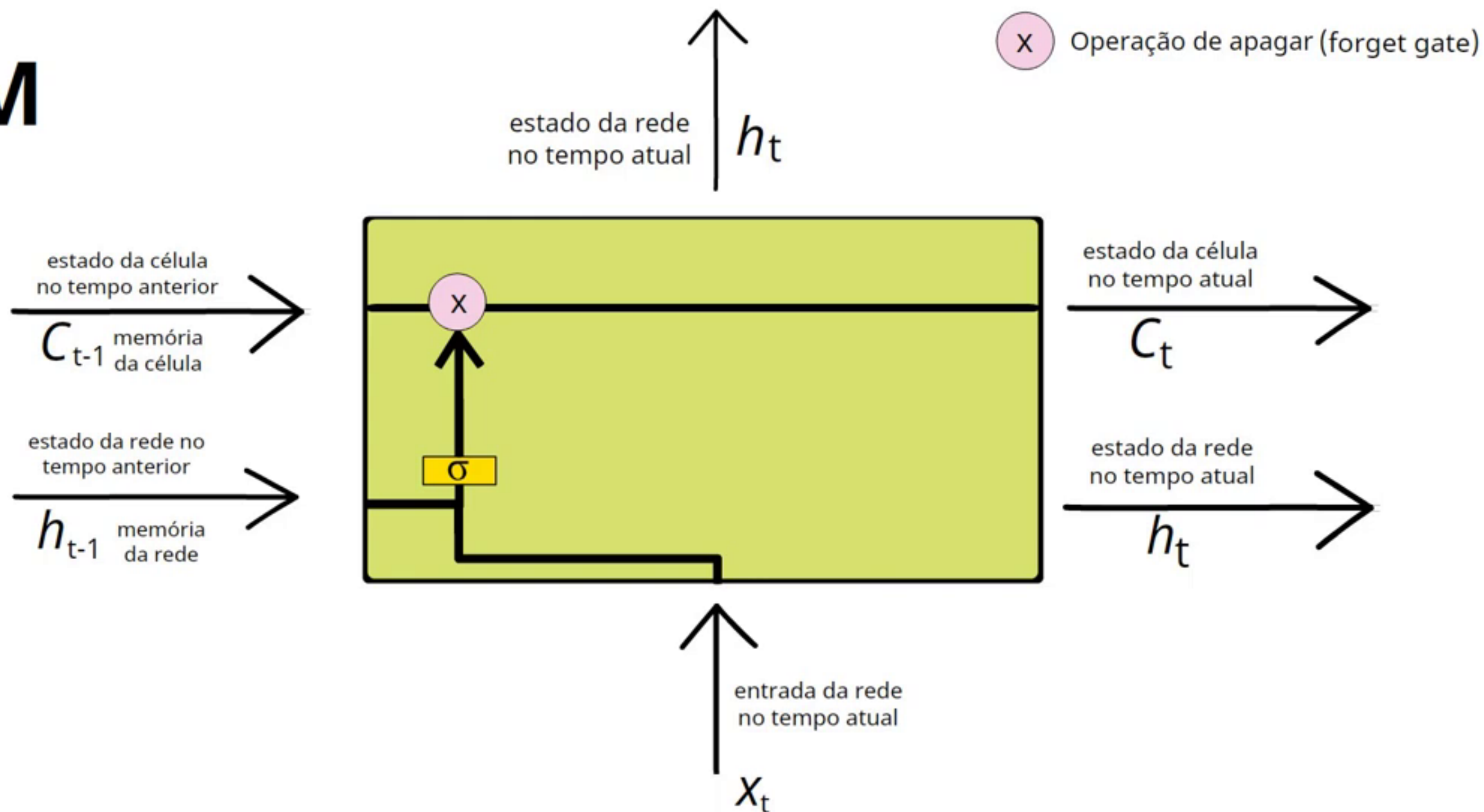
Quanto mais próximo de 1, mais informações serão preservadas
Se a sigmóide retornar zero, todas as informações serão descartadas

LSTM



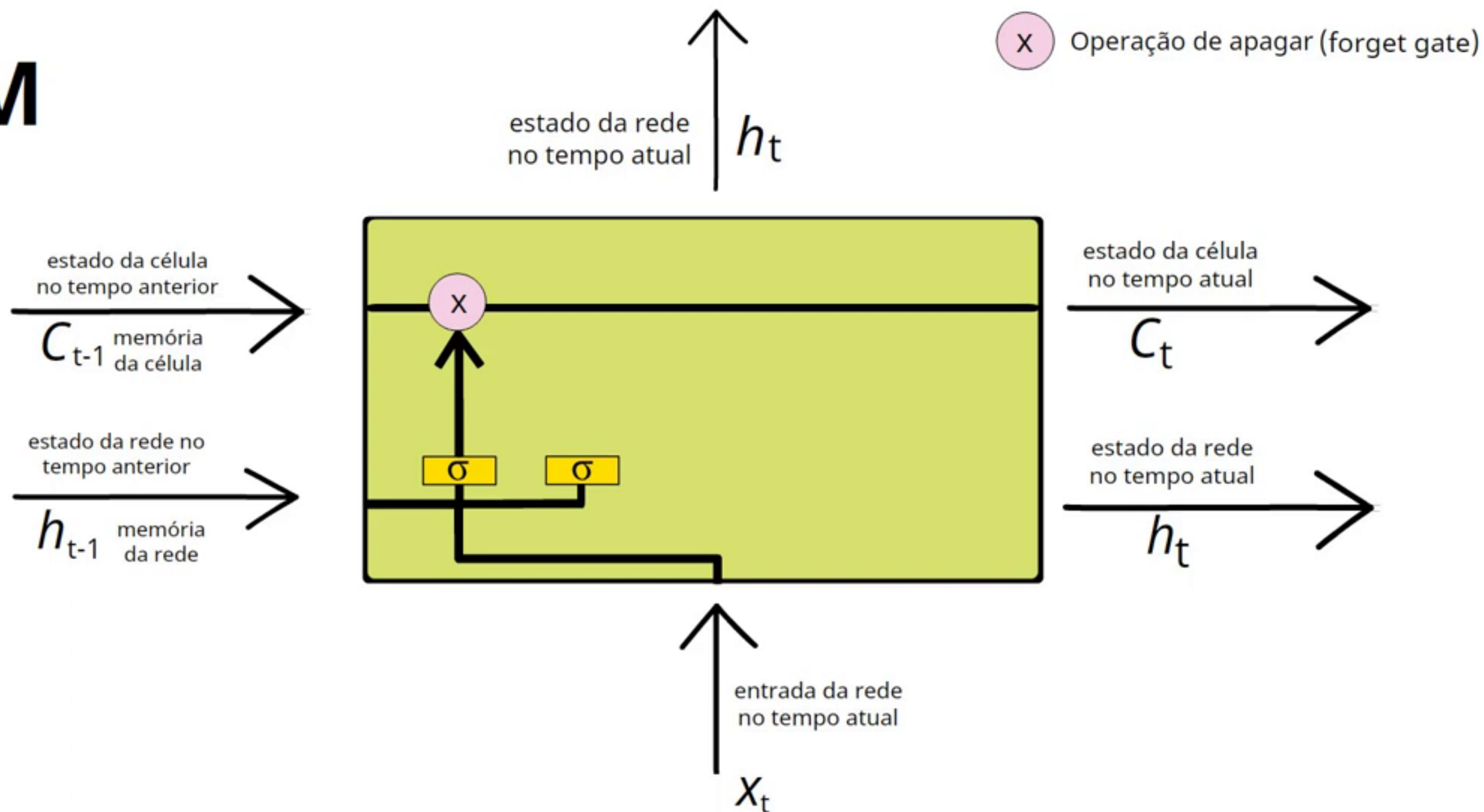
As informações que passaram pela função sigmóide são combinadas com o estado da memória no tempo anterior no "forget gate", que decide o que será apagado na memória

LSTM



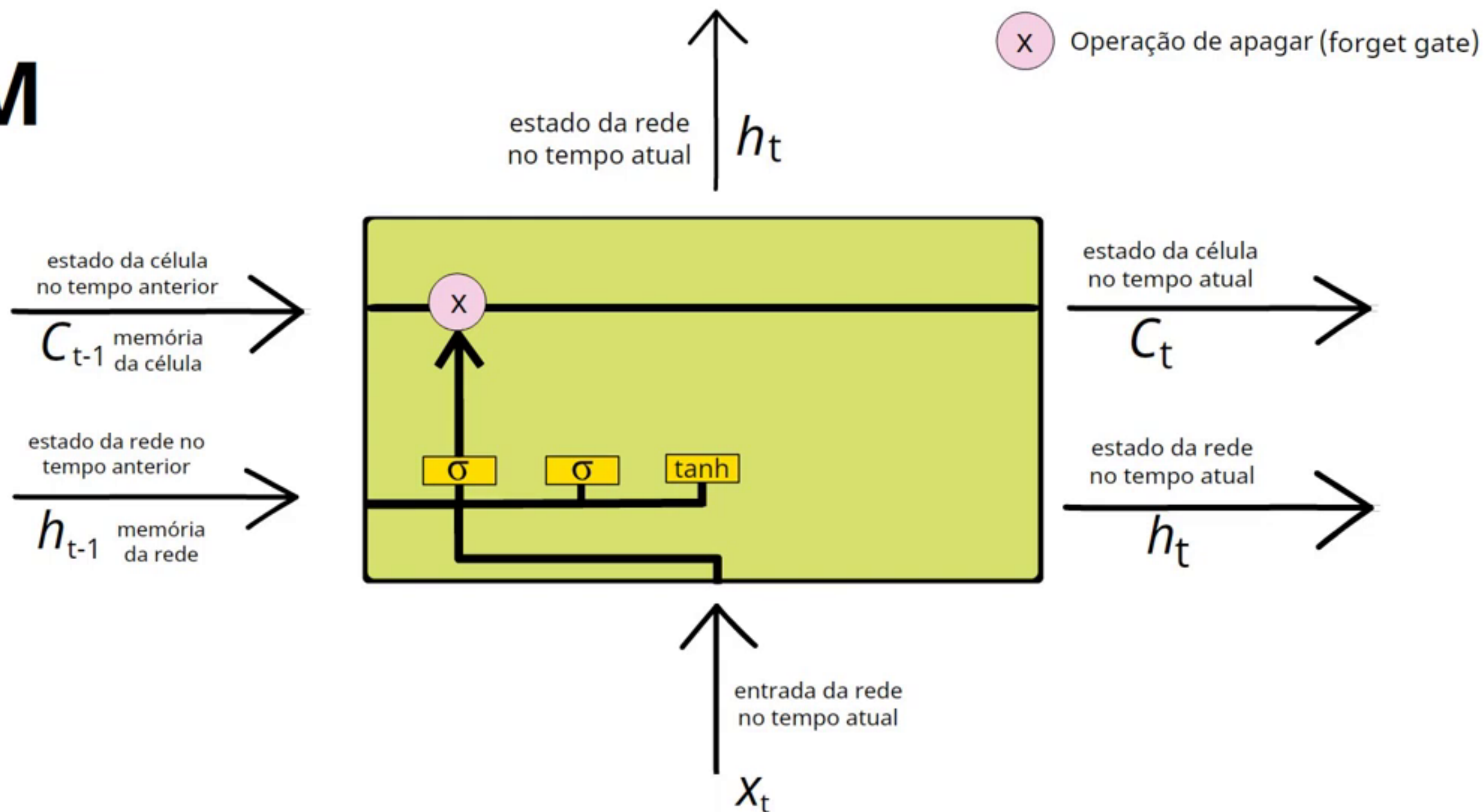
As informações que passaram pela função sigmóide são combinadas com o estado da memória no tempo anterior no "forget gate", que decide o que será apagado na memória

LSTM



Ainda com o objetivo de decidir quais informações novas serão armazenadas no estado da célula, as informações do estado anterior e da entrada no tempo atual também servem de entrada para uma função tangente hiperbólica

LSTM

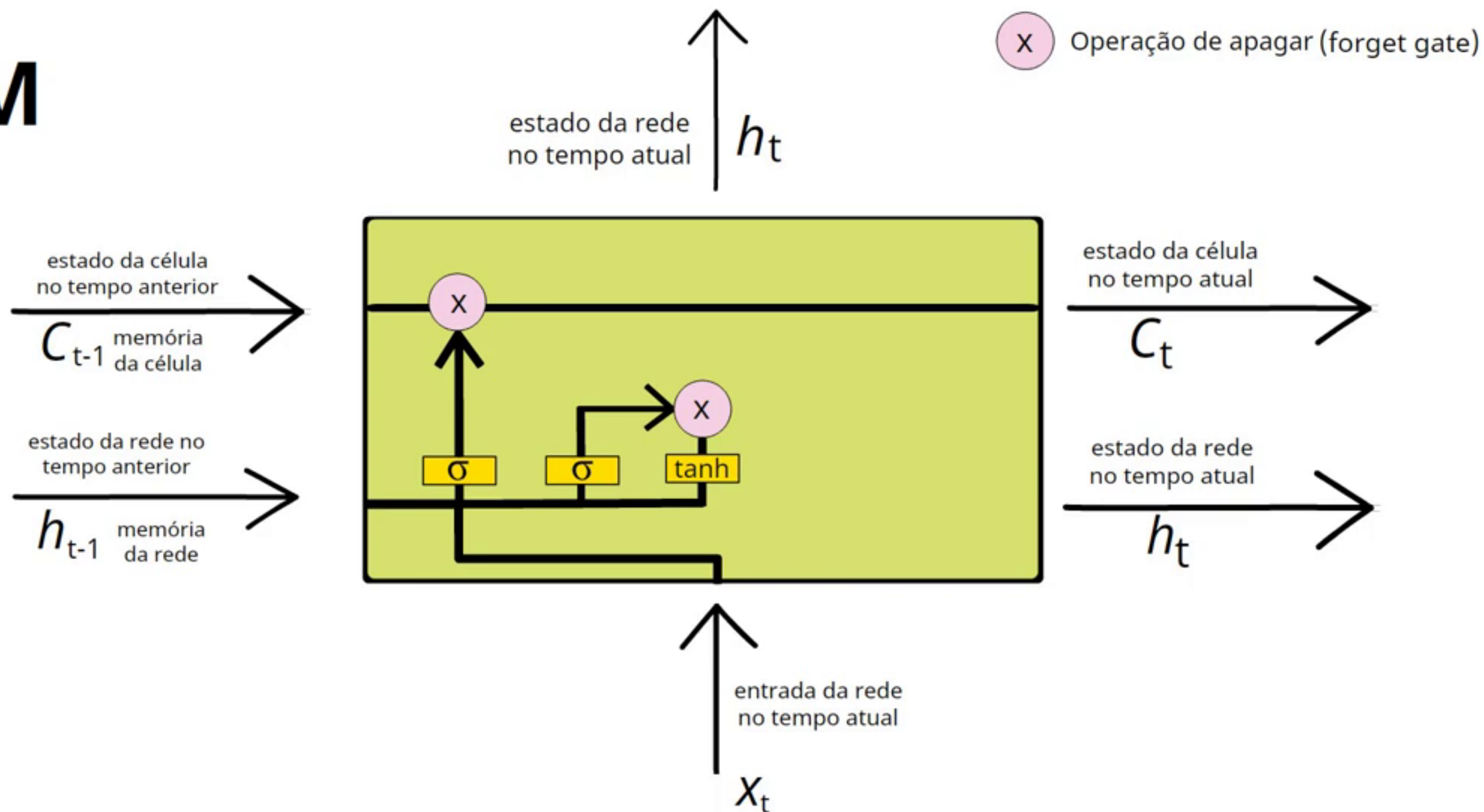


tanh Normaliza os valores para o intervalo $[-1, 1]$

Estabiliza a rede - trabalha com valores menores

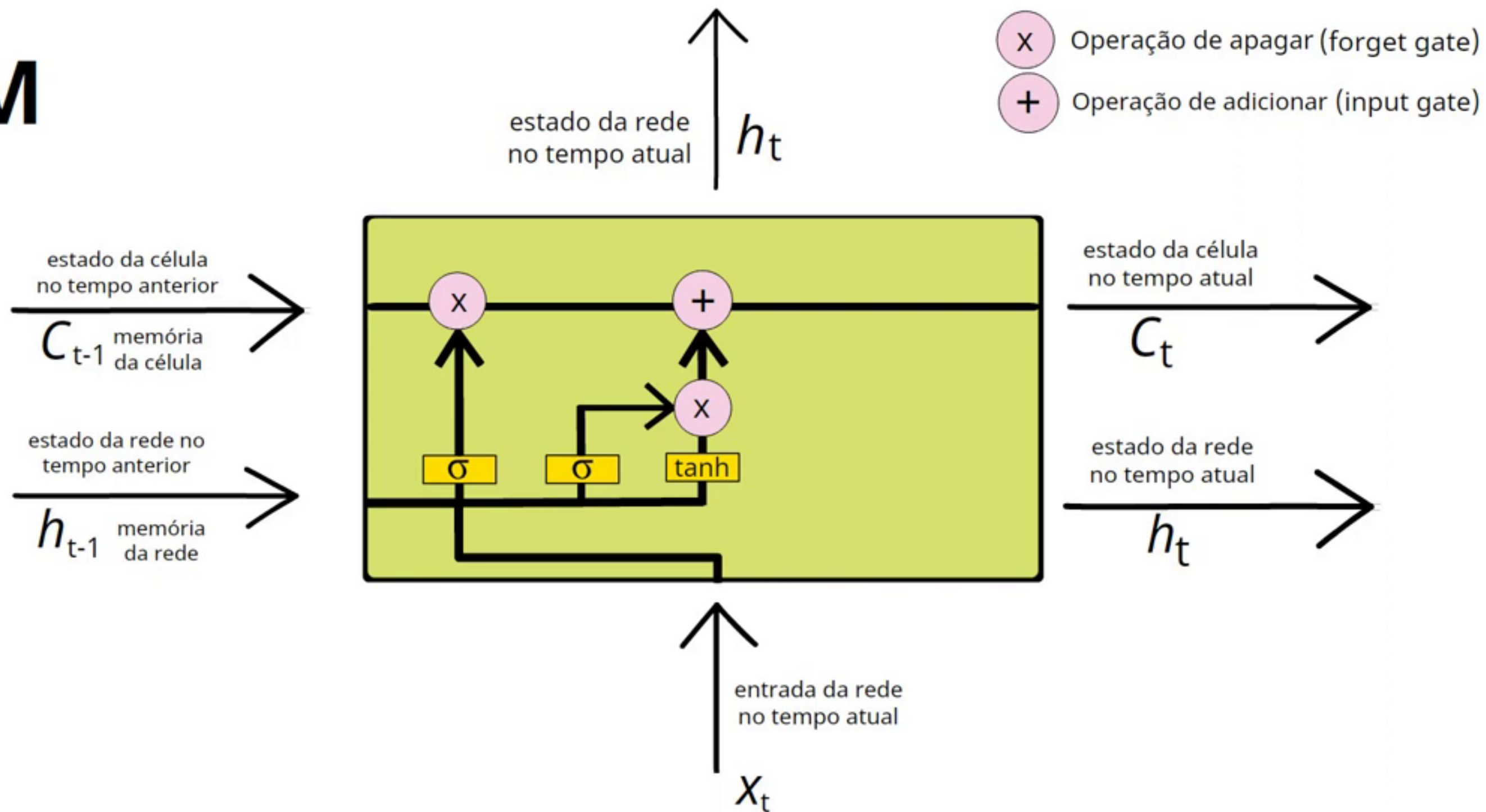
Introduz não linearidade - essencial para a rede aprender relações não triviais

LSTM



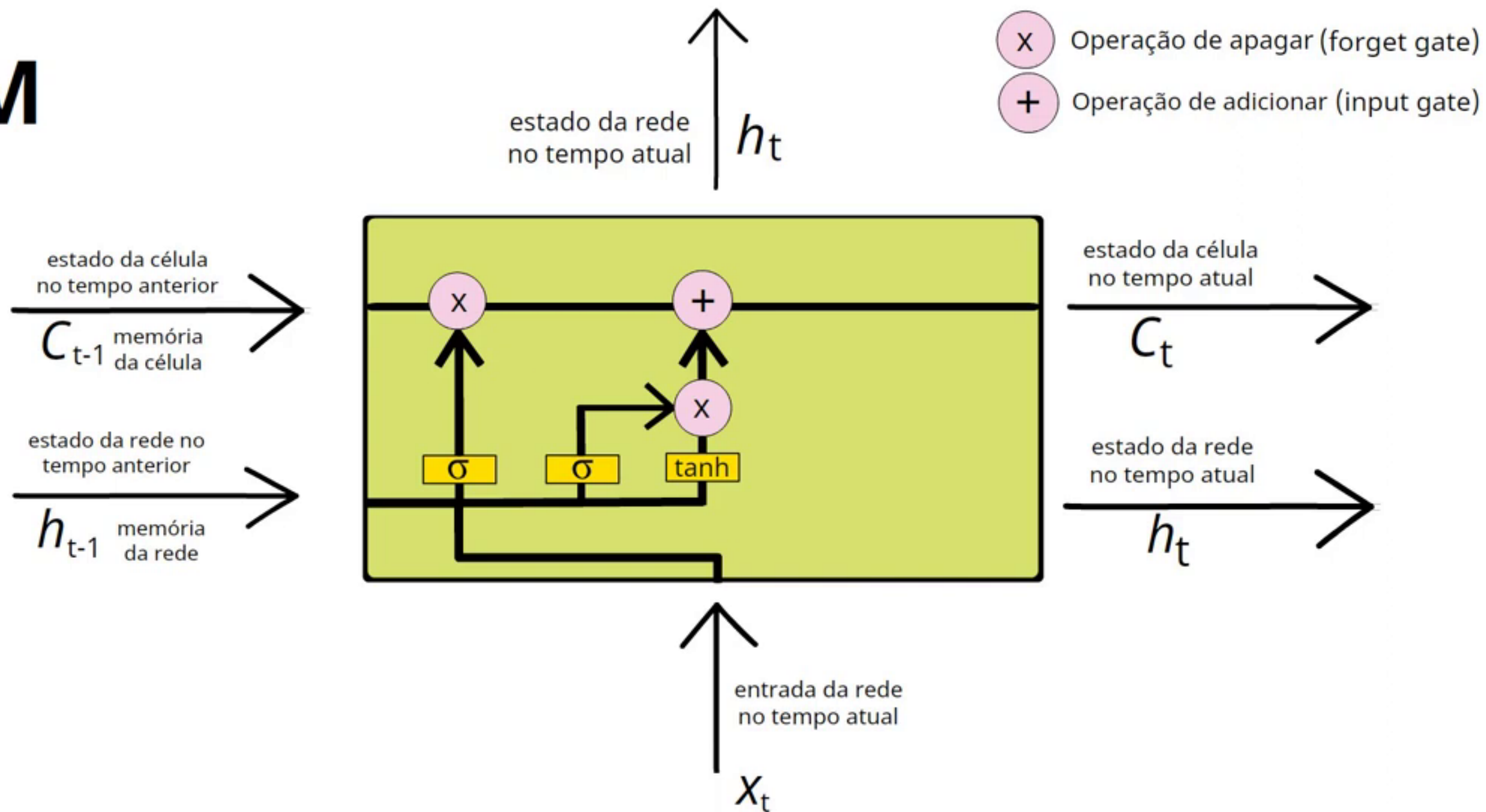
Depois de passar pelo "forget gate", a informação é combinada com a saída do "forget gate" anterior. Essa combinação é realizada por um "input gate", que decide quais dessas novas informações serão adicionadas à memória da rede

LSTM



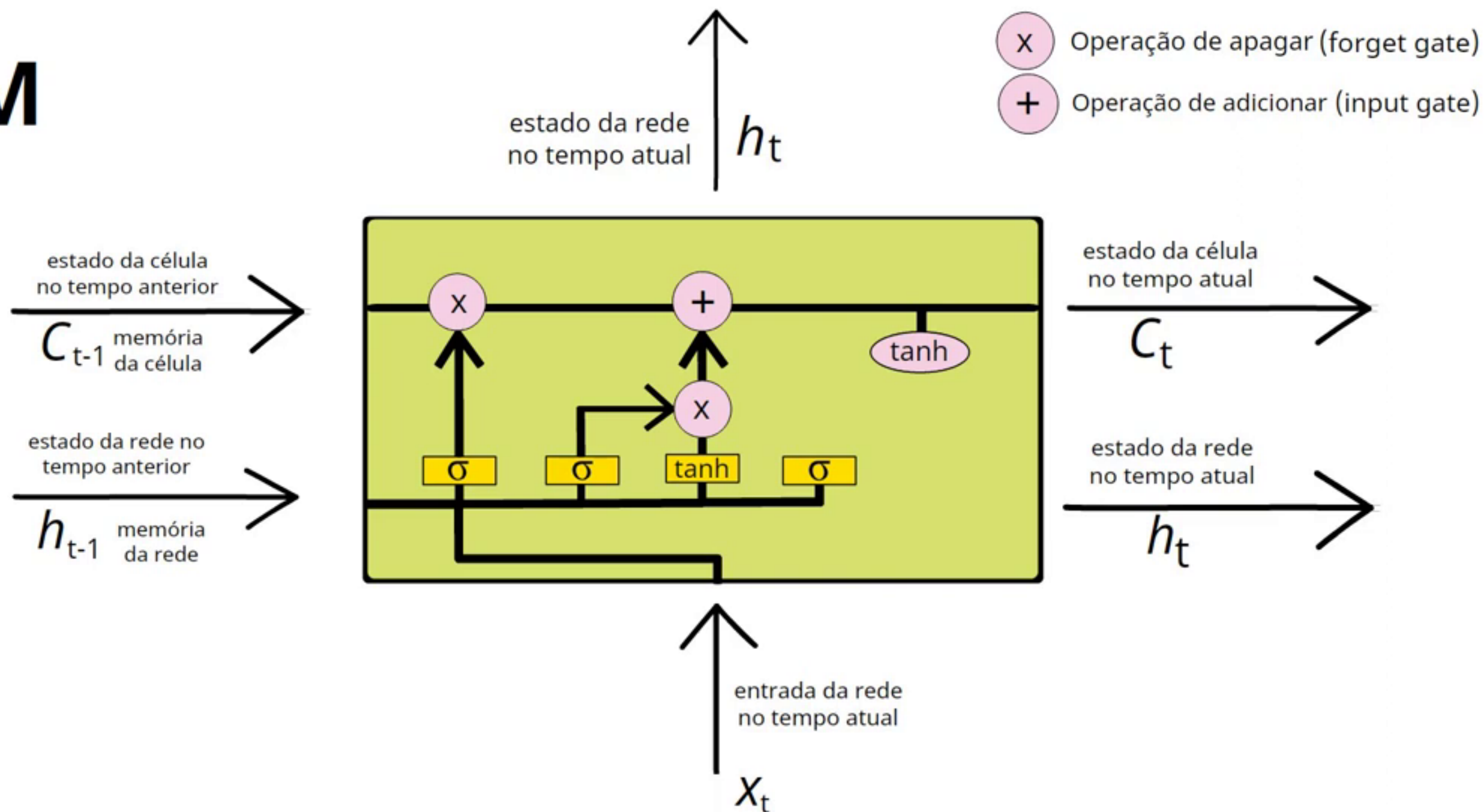
Neste momento, a rede já calculou o estado atual da célula de memória

LSTM



O próximo passo é decidir quais informações serão passadas adiante por meio do estado da rede

LSTM

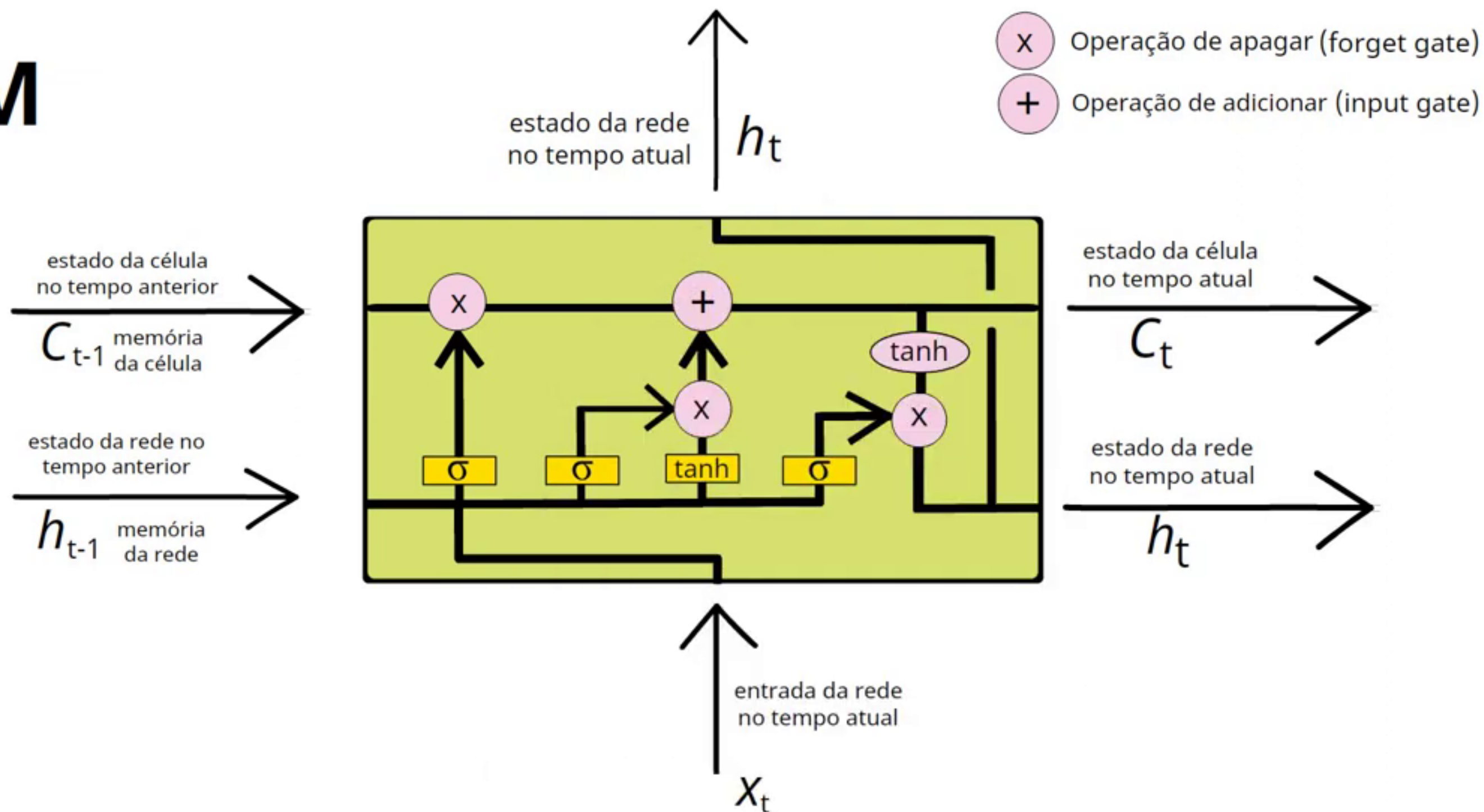


Primeiro, uma função tangente hiperbólica recebe o estado atual da célula

Uma sigmóide recebe informações da entrada atual e do estado anterior

A saída da sigmóide é combinada com a da hiperbólica em um "forget gate"

LSTM



Primeiro, uma função tangente hiperbólica recebe o estado atual da célula

Uma sigmóide recebe informações da entrada atual e do estado anterior

A saída da sigmóide é combinada com a da hiperbólica em um “forget gate”

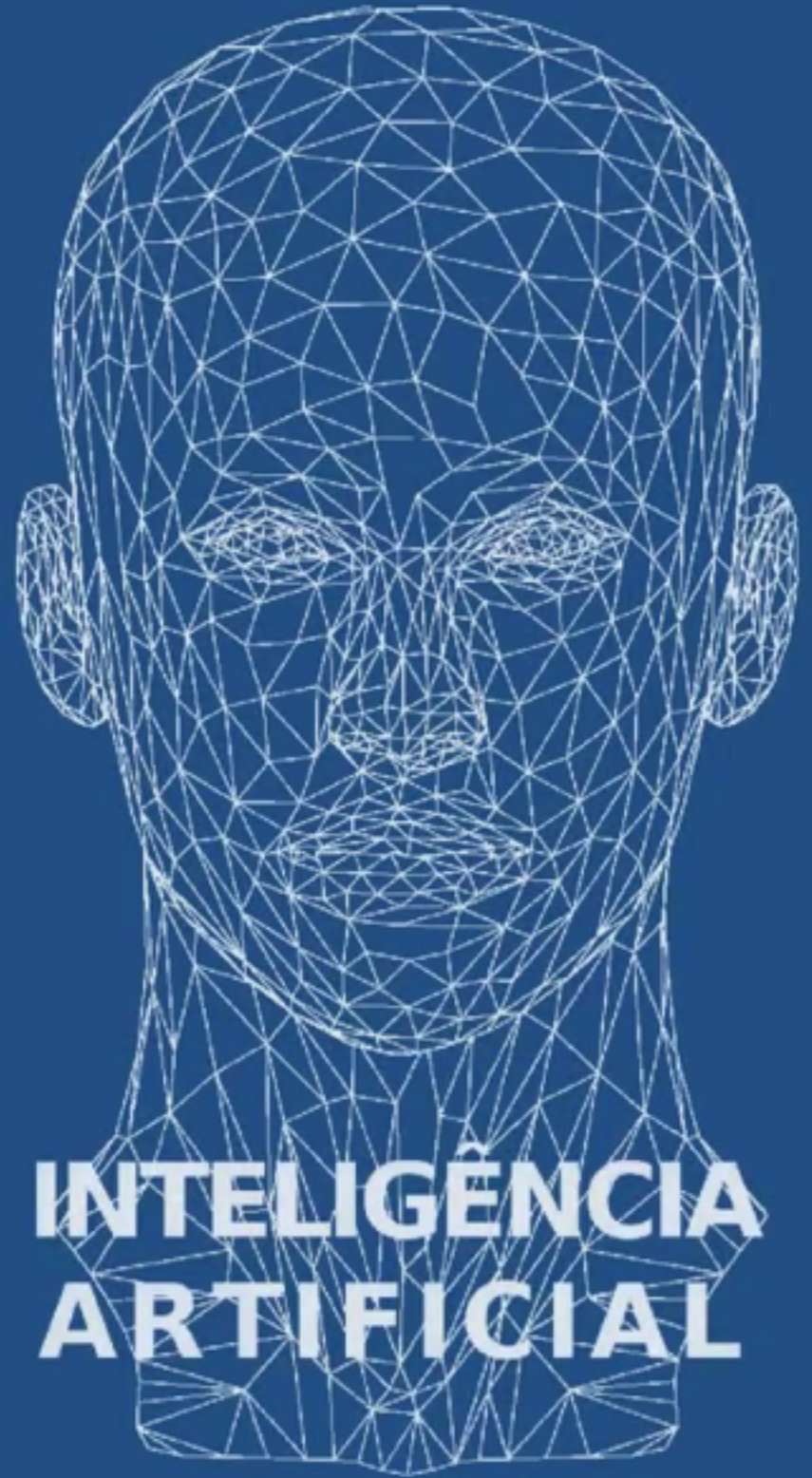
As informações resultantes são então enviadas para a saída e para a própria rede



Dissipação do gradiente

Vanishing Gradient





**INTELIGÊNCIA
ARTIFICIAL**