

Redes neurais recorrentes

Jones Granatyr



Introdução

- Dados sequenciais
- Prever a próxima ação (frames anteriores podem prever o frame atual)
- Processamento de linguagem natural
 - Previsão da próxima palavra em um texto
 - Tradução automática
 - Geração de poemas
- Geração de legendas em vídeos
- Séries temporais (time series)
 - Preço de ações na bolsa de valores
 - Temperatura
 - Crescimento populacional
 - Nível de poluição

Eu sou do Brasil
.....
.....
.....
..... eu falo português

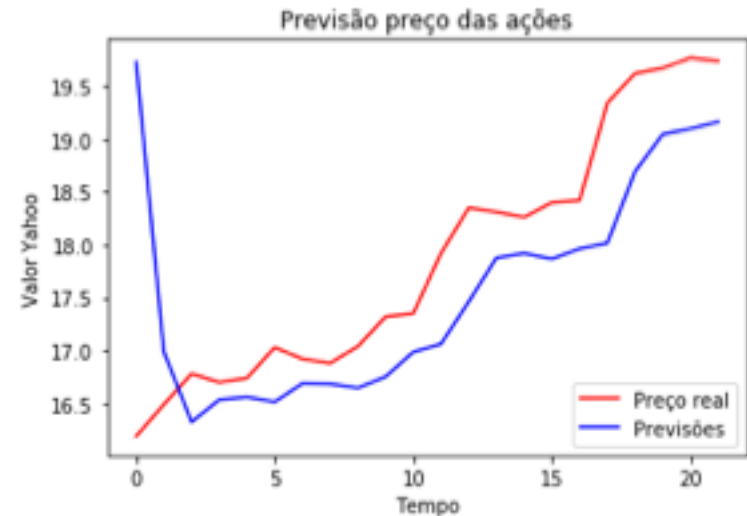
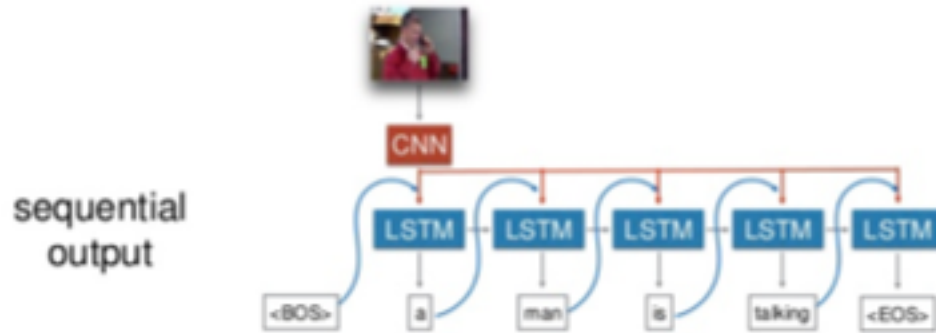
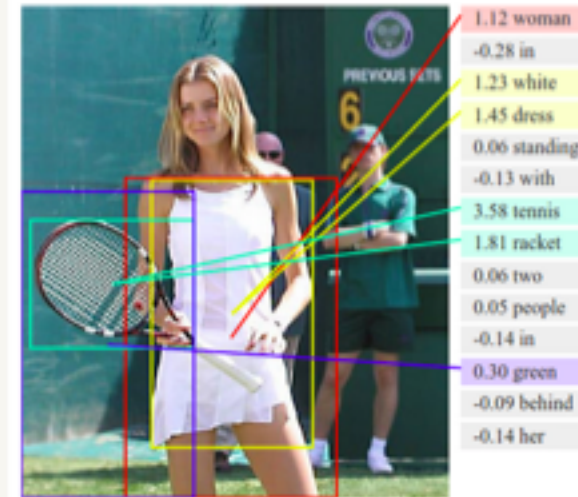
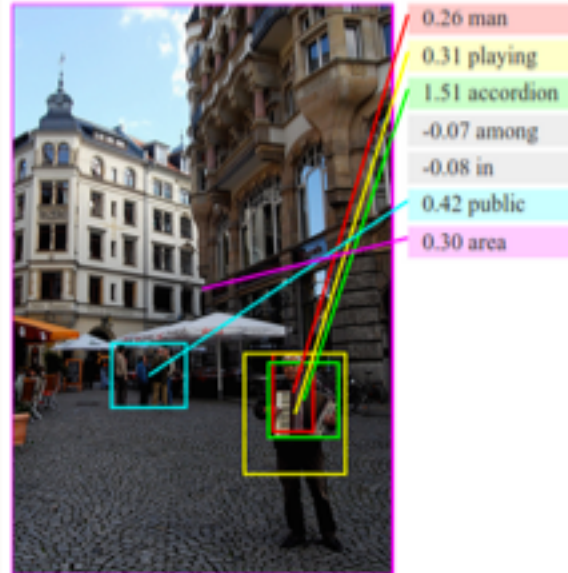
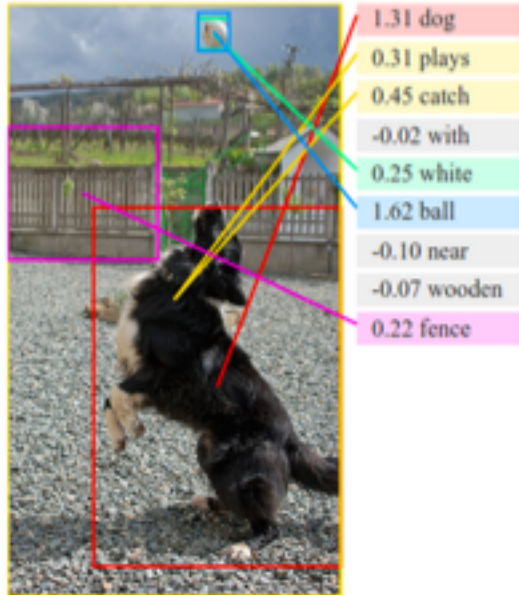


Image Description



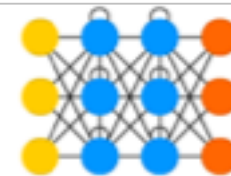
Fonte: <https://www.slideshare.net/ktoshik/kate-saenko-msr-russia-summer-school-modeling-images-video-text-caffe-di-part2>



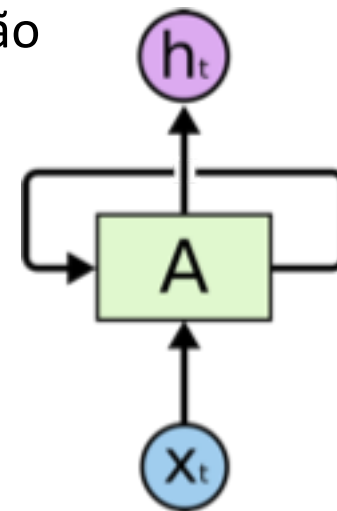
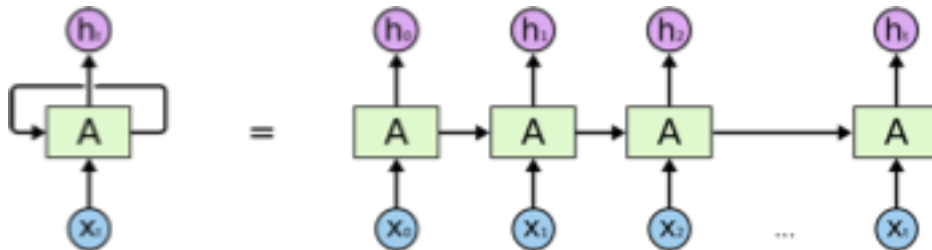
Fonte: <http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/>

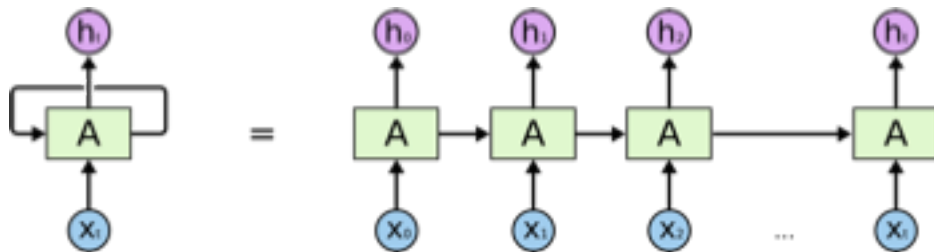
Parlamento Jovem Brasileiro

A lista dos estudantes pré-selecionados para a 15ª edição do Parlamento Jovem Brasileiro (PJB) foi divulgada pela Câmara Federal dos Deputados na segunda-feira, 25. Os 4 estudantes tocaninenses, selecionados pela banca da Secretaria de Estado da Educação, Juventude e Esportes (Seduc), são Ilan Rubins Carneiro de Carvalho, Louise Xavier Macedo e Thaissa Neres de Oliveira, da Escola Estadual Dr. Joaquim Pereira da Costa, de Gurupi, e Laicy Conceição Rodrigues Soares, do Colégio Estadual Dr. Abner Araújo Pacini, do município Almas.



- Para entender o final de uma frase você precisa saber o que foi dito antes
- Redes neurais tradicionais não armazenam informações no tempo (previsões independentes)
- Redes neurais com loops que permitem que a informação persista
- Múltiplas cópias de si mesmas





Fonte: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

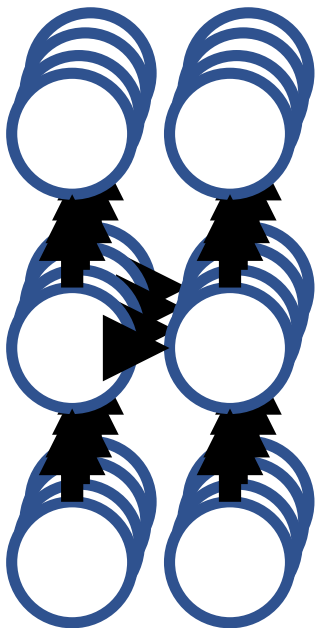
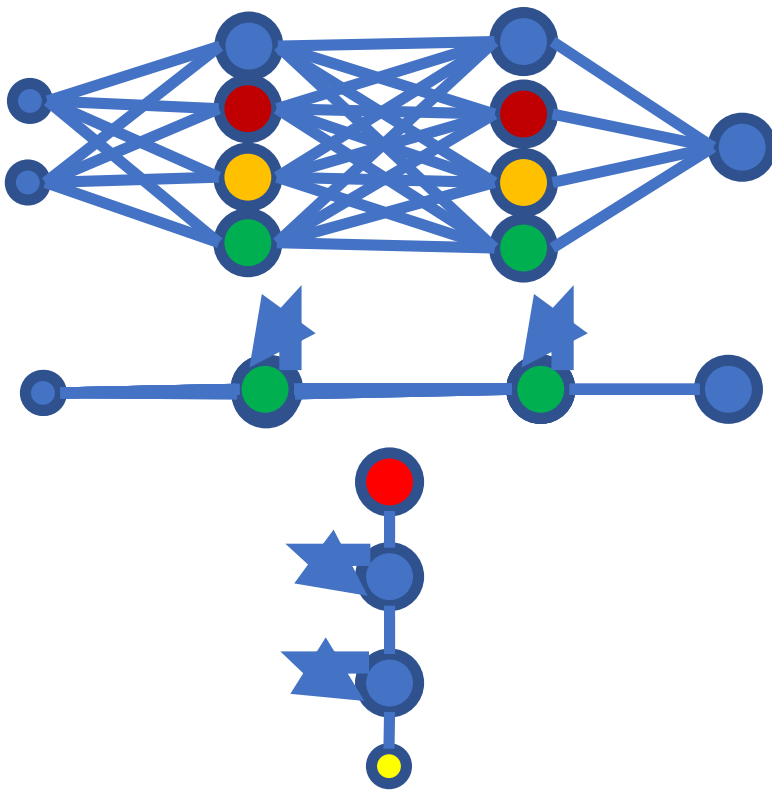
Espremer as camadas e cada camada da RNN será somente um único nó, mas cada nó representa uma camada que pode conter vários neurônios. A camada escondida é representada por somente um único neurônio que representa 4 neurônios

Camada escondida é conectada a ela mesma, mandando uma resposta para frente e também se realimentando

Backpropagation para treinar, desenrolando a rede neural para termos uma rede neural padrão

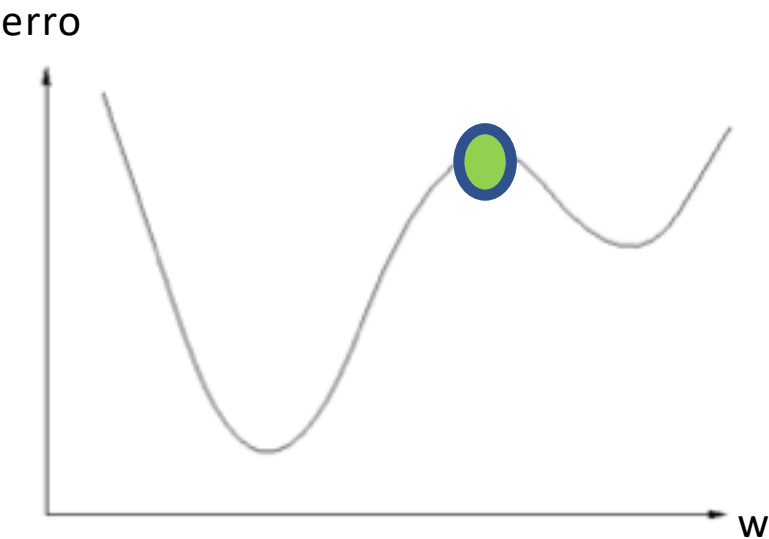
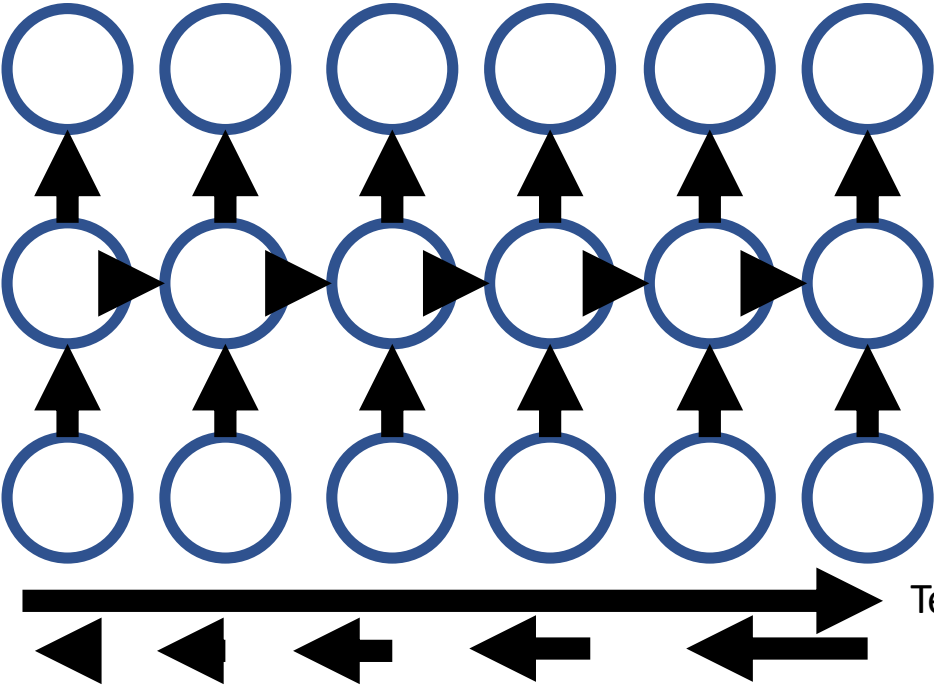
Os pesos são compartilhados, a memória é rápida, curta e lembra o que aconteceu apenas nas últimas interações

A auto ligação é o temporal no loop



Milhares de camadas

Os pesos são multiplicados várias vezes
e o resultado será cada vez menor que
não fará alteração



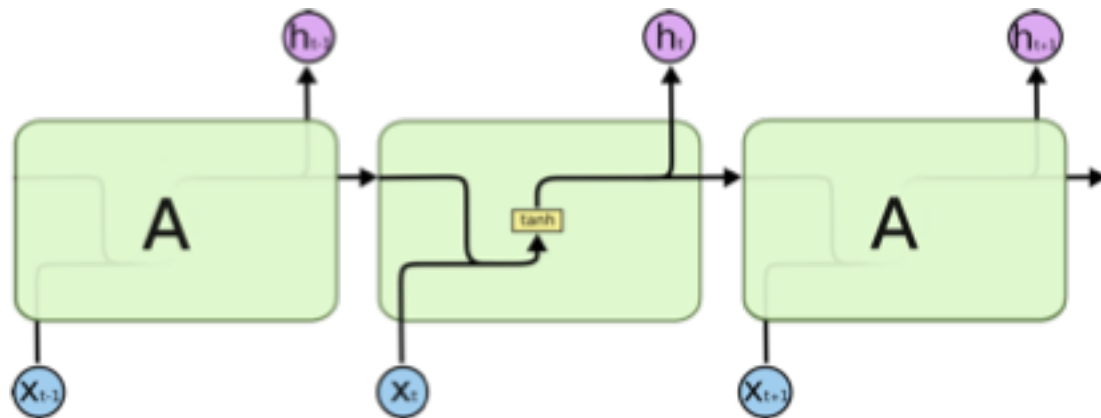
Gradiente desaparecendo (vanish gradient)

- Vanish
 - Backpropagation through time (BPTT): algoritmo semelhante que também fluirá para trás a partir do tempo futuro para os tempos atuais
 - Xavier initialization
 - LSTM (long short term memory)
- Exploding gradient: o valor poderá ficar muito grande
 - Não visitar todas as camadas ocultas
 - RMSProp: divide a taxa de aprendizagem por uma média exponencialmente decrescente do quadrado do gradiente
 - Clipping gradiente (grampo)

LSTM (Long-short term memory)

- “As nuvens estão no **céu**” (o espaço entre as palavras é pequeno, pode usar a informação passada)
- “Eu sou do Brasil...**eu falo português**” (precisa ter o contexto do Brasil)
- É difícil para uma rede neural recorrente “simples” aprender esse padrão, por isso existem as LSTM
- Adiciona células de memória na rede neural e manipula essas células
- Aprende dependências de longo prazo

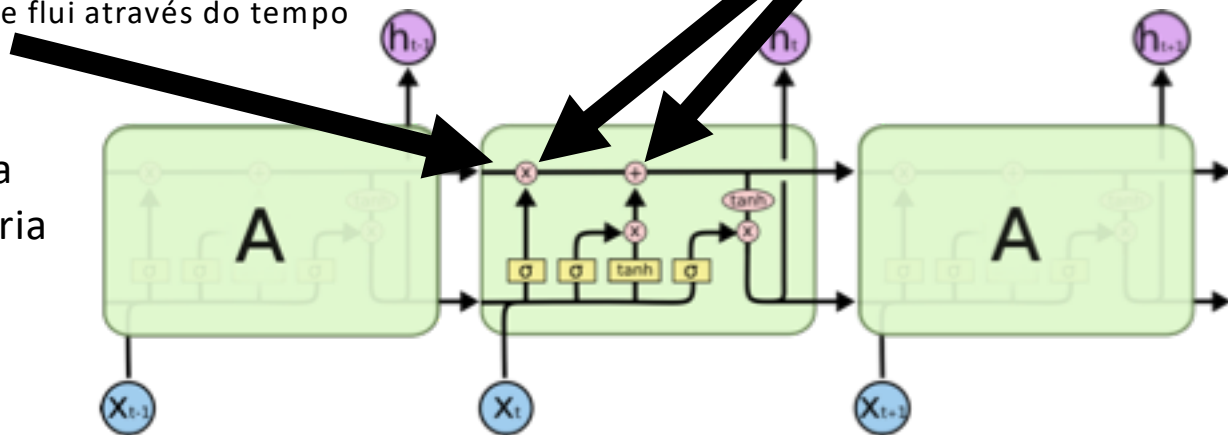
LSTM (Long-short term memory)



Operações de apagar e adicionar, o que evita os problema no gradiente

Célula de memória que flui através do tempo

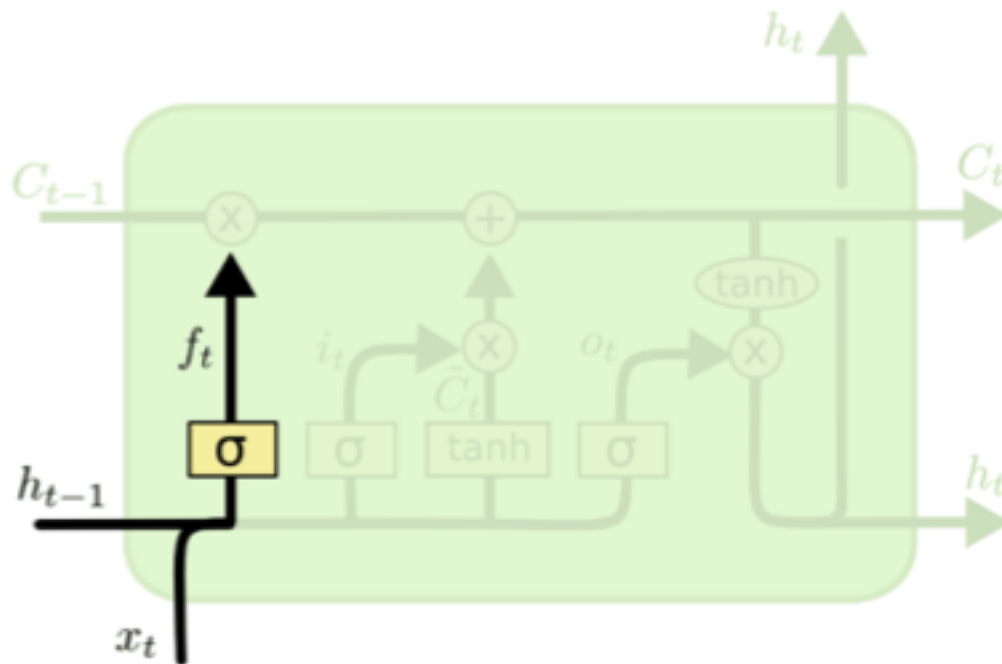
Forget gate: liberar da memória
Input gate: adicionar na memória
Output gate: ler da memória



1. Decidir o que será apagado

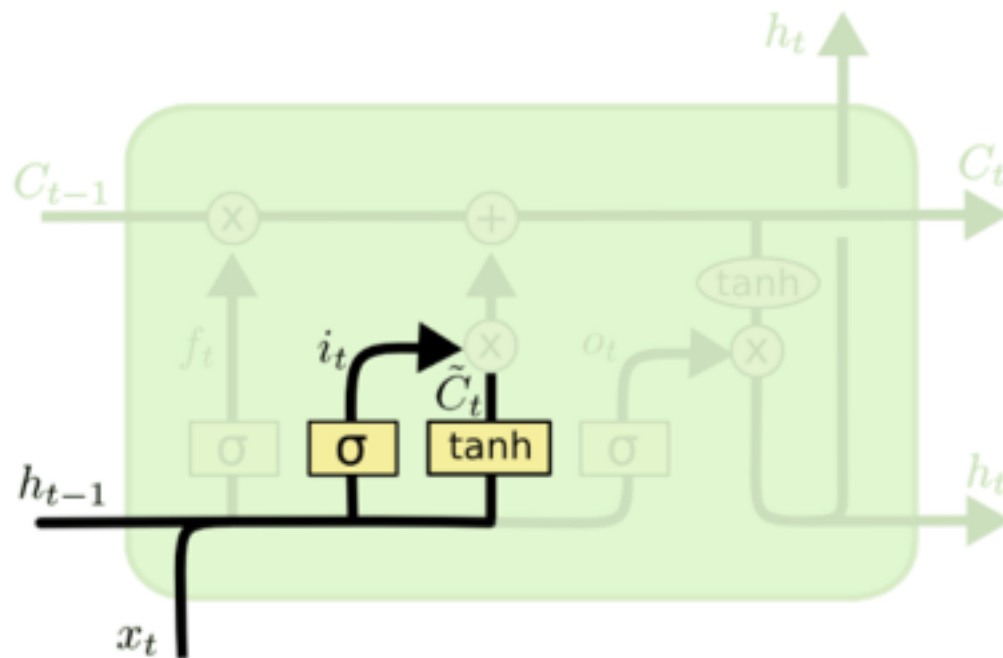
Função sigmoide porque retornará valores entre 0 e 1

Se o valor for 0 o dado não é importante e é apagado da memória



A memória pode incluir o gênero da pessoa para que os pronomes corretos possam ser usados. Quando a rede encontra uma nova pessoa, pode apagar o gênero da pessoa anterior

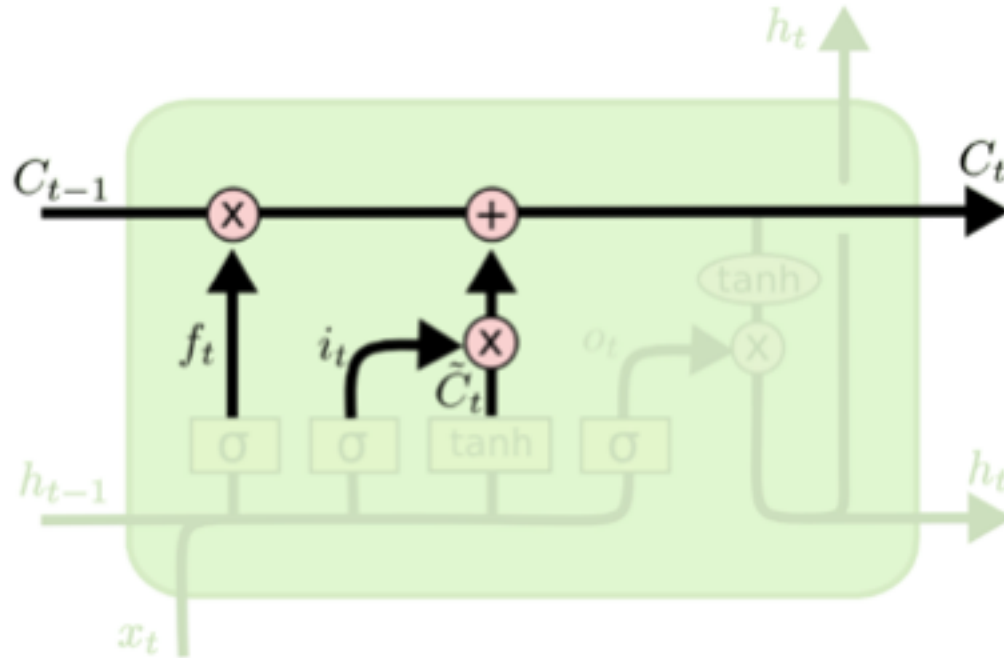
2. Decidir o que será armazenado



Decidir quais valores serão alterados e a função tangente hiperbólica cria um vetor de dos novos candidatos

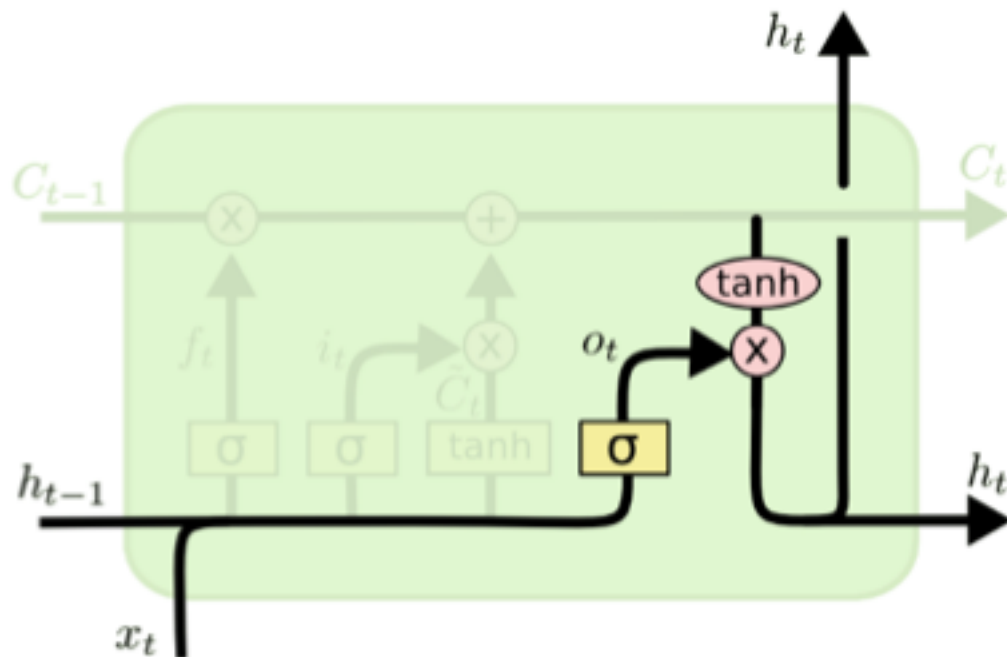
Adicionar o novo gênero da pessoa que foi apagado anteriormente

3. Atualizar o estado antigo



As etapas anteriores decidiram o que apagar e o que armazenar, e agora essas etapas são executadas

4. Decidir qual será a saída



Aplica a função sigmoide e coloca os valores entre -1 e 1 para retornar somente as partes necessárias (se encontrou uma pessoa/objeto, pode retornar se é no plural ou singular)

Conclusão

