

Redes Neurais Recorrentes

Computação Natural
Gisele Pappa

Introdução

- Redes Neurais Recorrentes (RNN) são uma família de redes própria para representar dados de sequências

Em que direção ele
vai seguir agora?



Introdução

- Redes Neurais Recorrentes (RNN) são uma família de redes própria para representar dados de sequências



Dados de Sequência

Reconhecimento de
fala



“The book is on the table.”

Geração de música



Análise de
Sentimentos

“There is nothing to
like in this movie.”



Análise de
Sequências de DNA

AGCCCCTGTGAGGAA
CTAG



AG**CCCCTGTGAGGAA**
CTAG

Plasmid sequence AP4 Oligo

Dados de Sequência

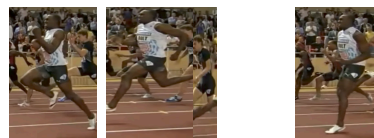
Tradução
automática

Voulez-vous chanter
avec moi?



Do you want to sing
with me?

Reconhecimento de
ações em vídeos



Running

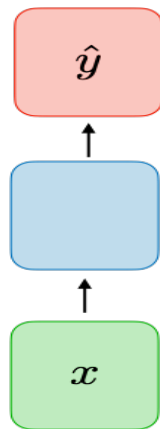
Reconhecimento de
Entidades nomeadas

Yesterday, Harry
Potter met Hermione
Granger.



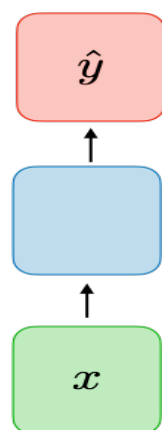
Yesterday, **Harry
Potter** met **Hermione
Granger**.

Modelagem de Sequências

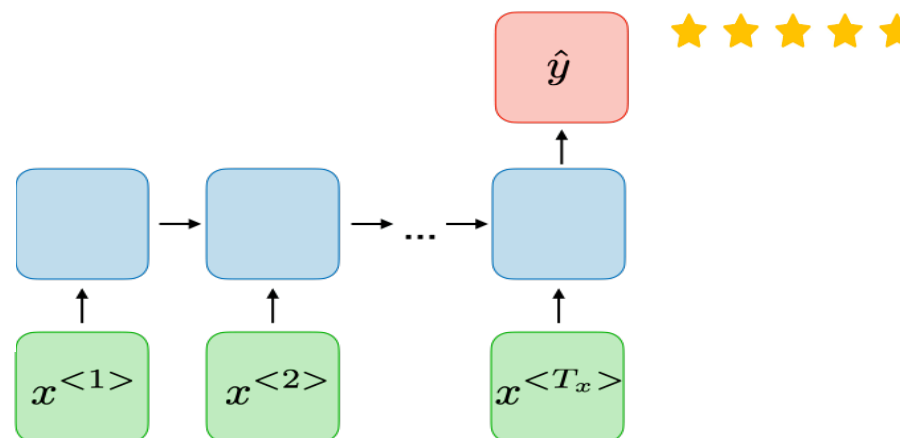


Um para um
Classificação
(Prever o gênero de
uma música)

Modelagem de Sequências



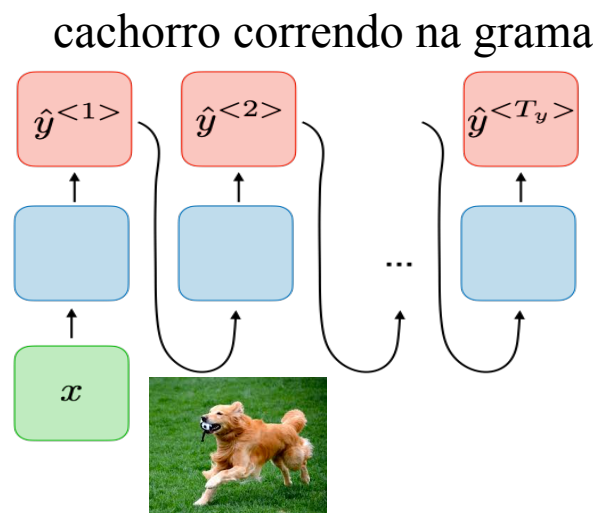
Um para um
Classificação
(Prever o gênero de
uma música)



There is nothing to like in this movie

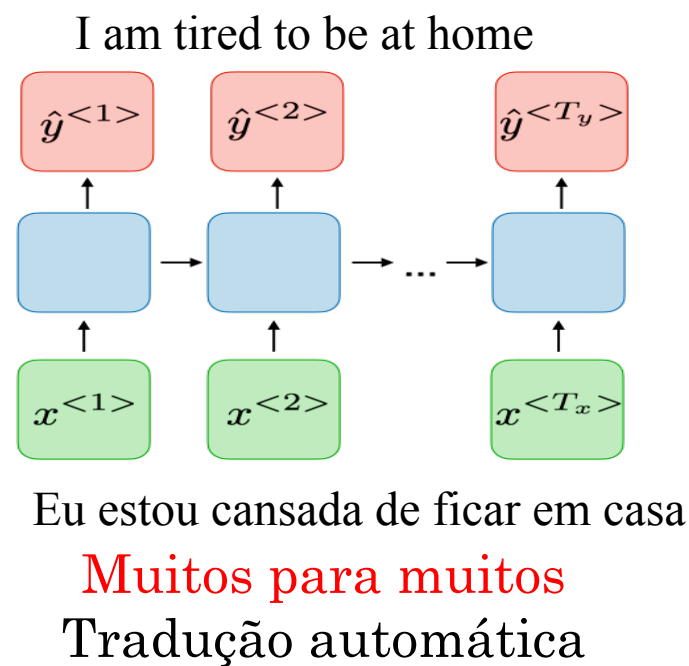
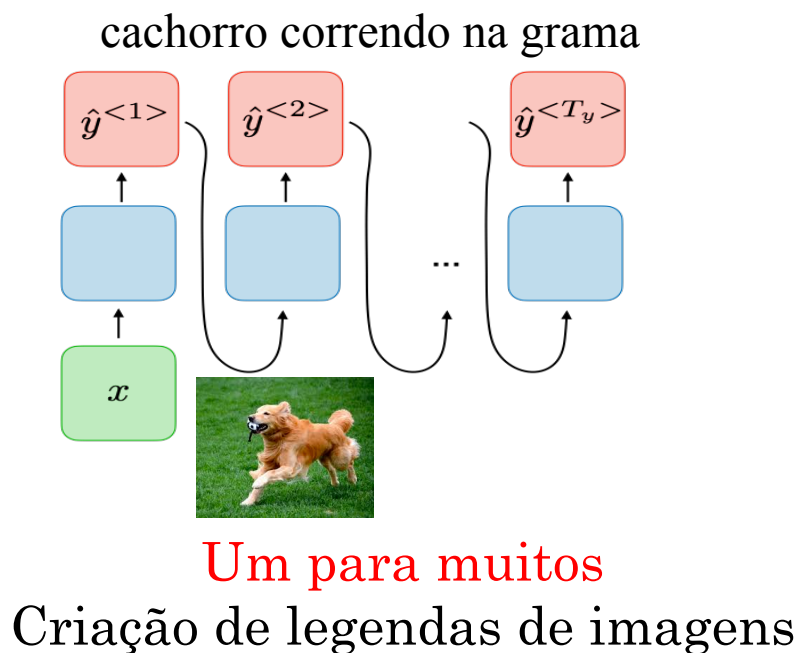
Muitos para um
Análise de Sentimentos

Modelagem de Sequências

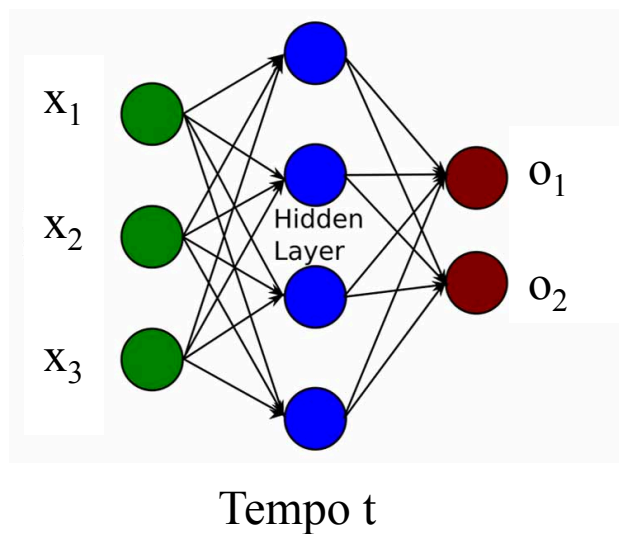


Um para muitos
Criação de legendas de imagens

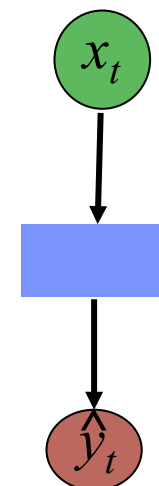
Modelagem de Sequências



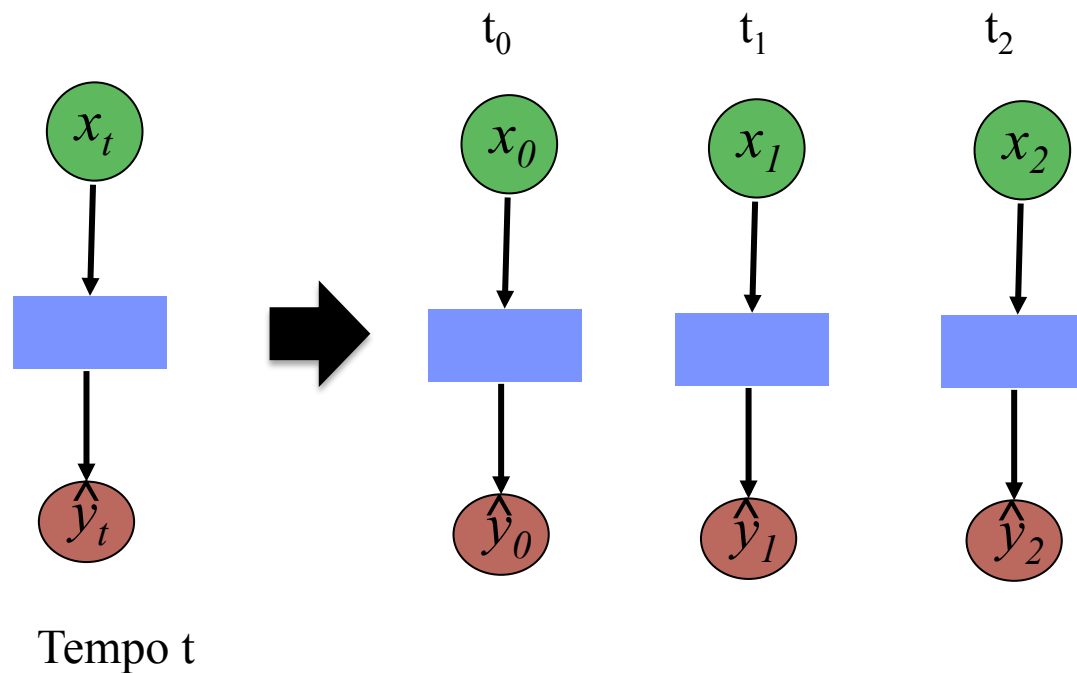
Feed Forward vs Sequências



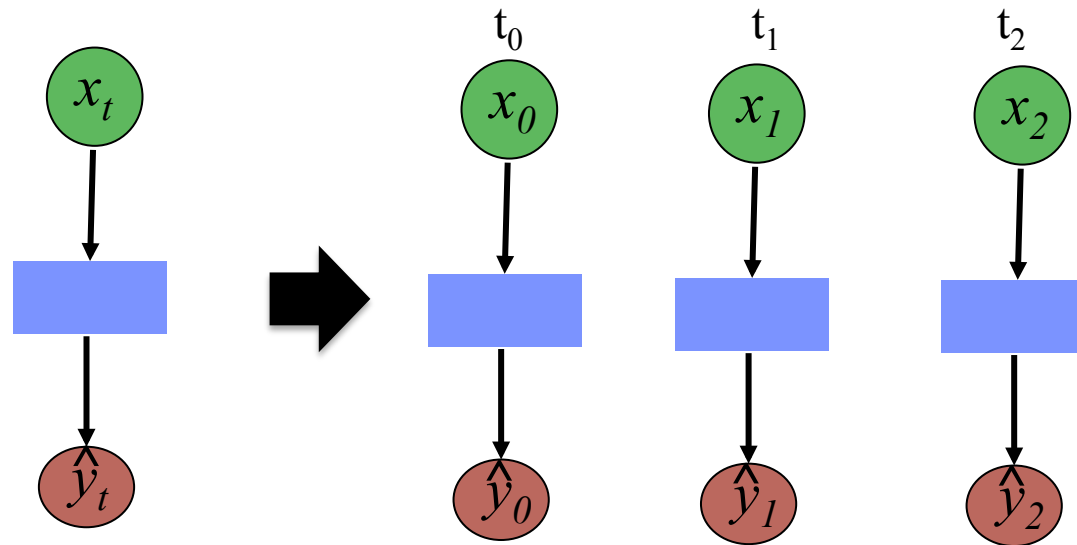
$$\hat{y}_t = o_t$$



Feed Forward vs Sequências



Feed Forward vs Sequências



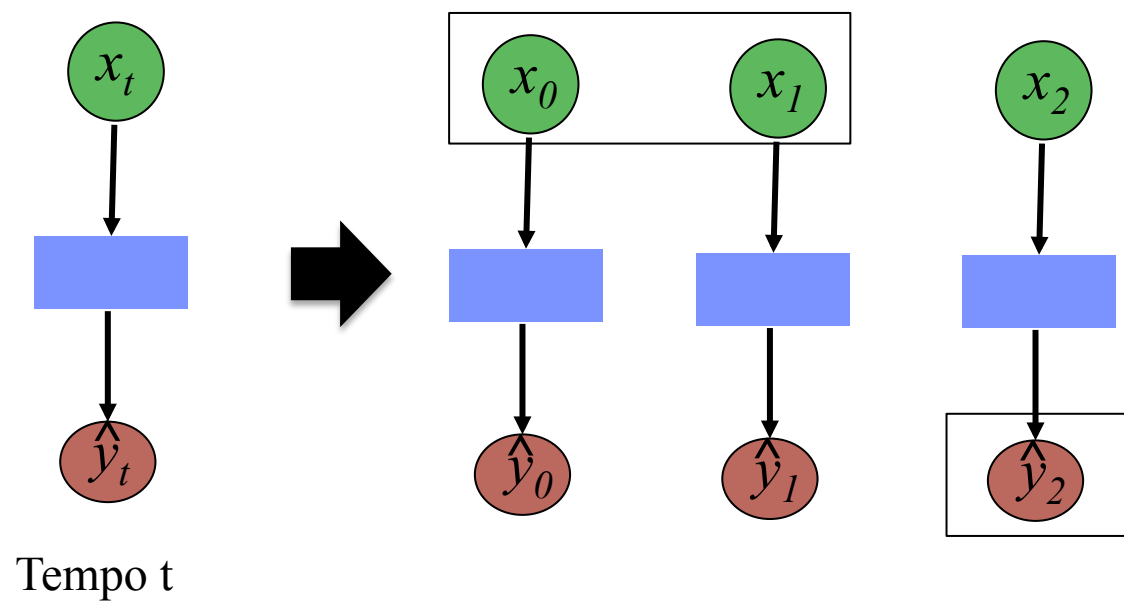
Tempo t

- Rede teria parâmetros distintos para cada entrada
- Rede teria que aprender todas as regras da linguagem separadamente para cada posição da sentença

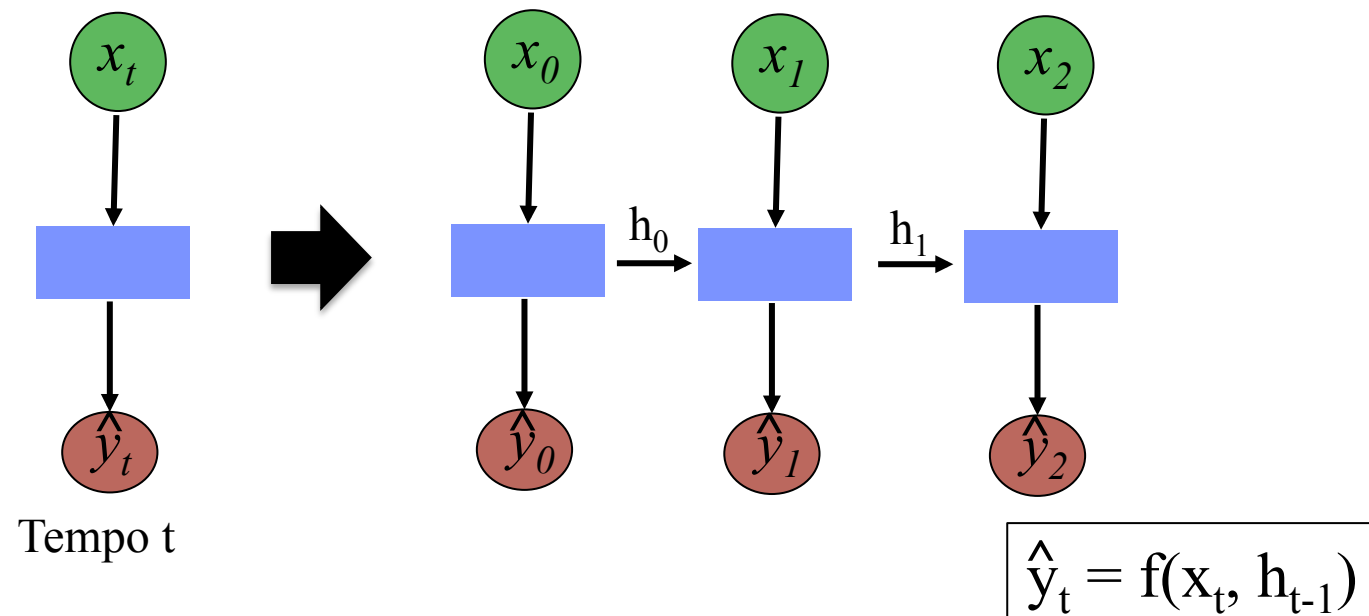
Feed Forward vs Sequências

- Considere 2 sentenças:
Eu entrei na universidade em 2015.
Em 2015, eu entrei na universidade.
- Se o objetivo é que a rede identifique o ano que você entrou na universidade, ela deve reconhecer o ano de 2015 sendo ele a 6ª ou 2ª palavra da sentença.

Feed Forward vs Sequências

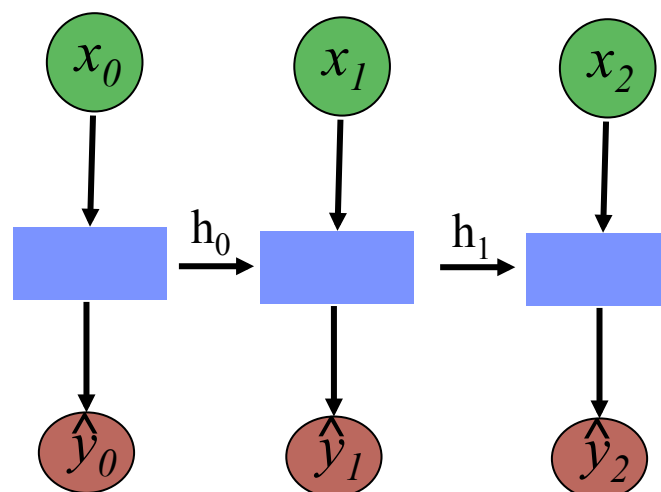
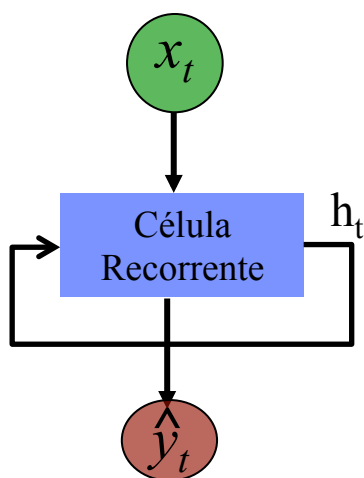


Redes Neurais Recorrentes (RNN)



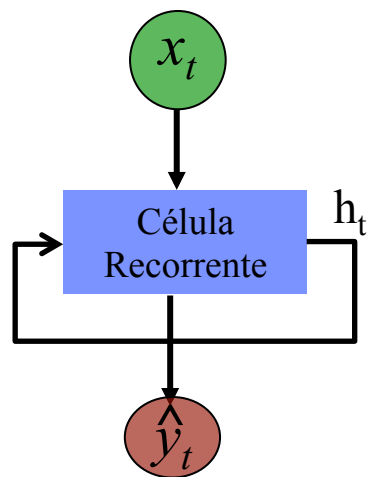
- h corresponde ao estado/memória da rede, e é atualizado a cada unidade de tempo

Redes Neurais Recorrentes (RNN)



$$\hat{y}_t = f(x_t, h_{t-1})$$

Redes Neurais Recorrentes (RNN)



Função com pesos W Estado anterior

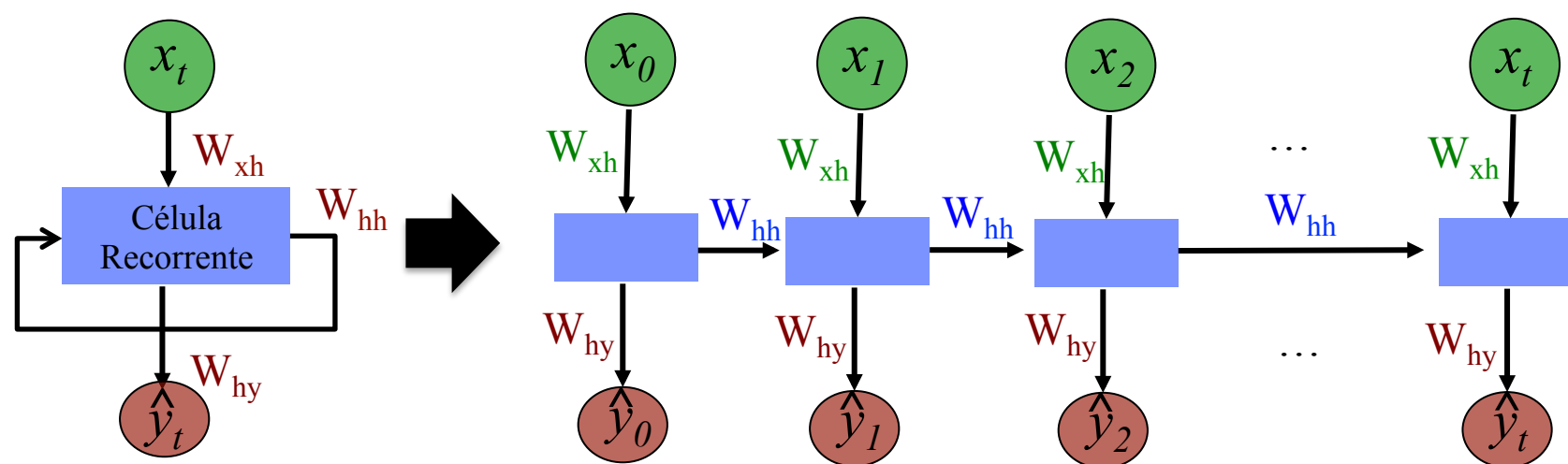
$$h_t = f_W(x_t, h_{t-1})$$

Estado da célula entrada

A mesma função e conjunto de parâmetros são utilizados para cada unidade de tempo da sequência

- h_t corresponde ao estado/memória da rede, e é atualizado a cada unidade de tempo

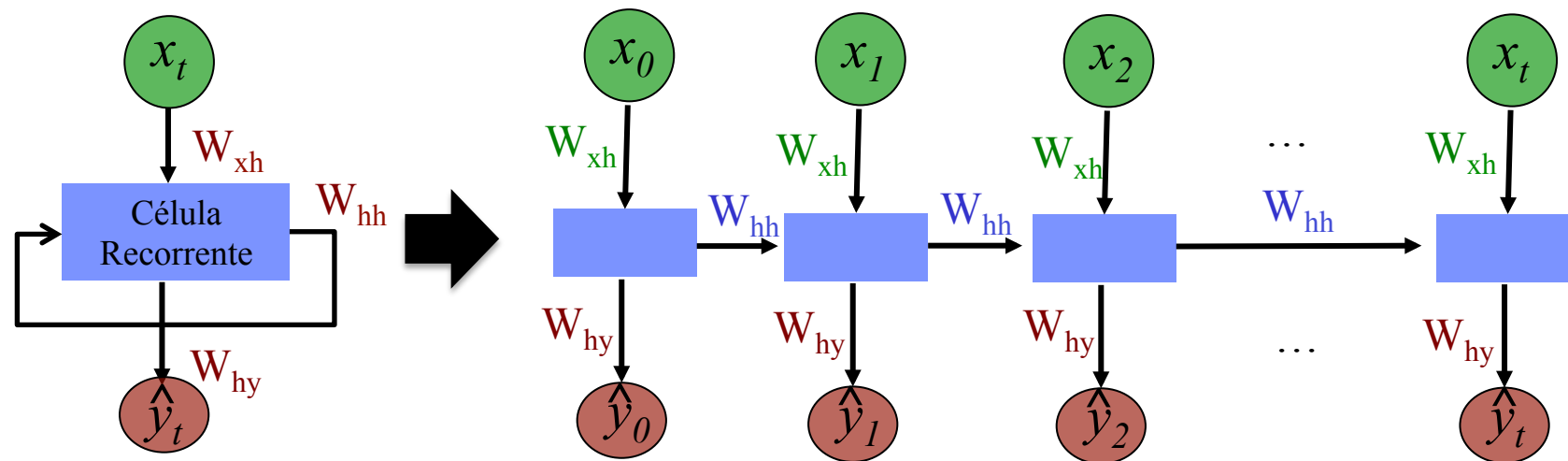
Redes Neurais Recorrentes (RNN)



Novamente, note que as matrizes de pesos aprendidas são **compartilhadas** por todos os períodos de tempo

Redes Neurais Recorrentes (RNN)

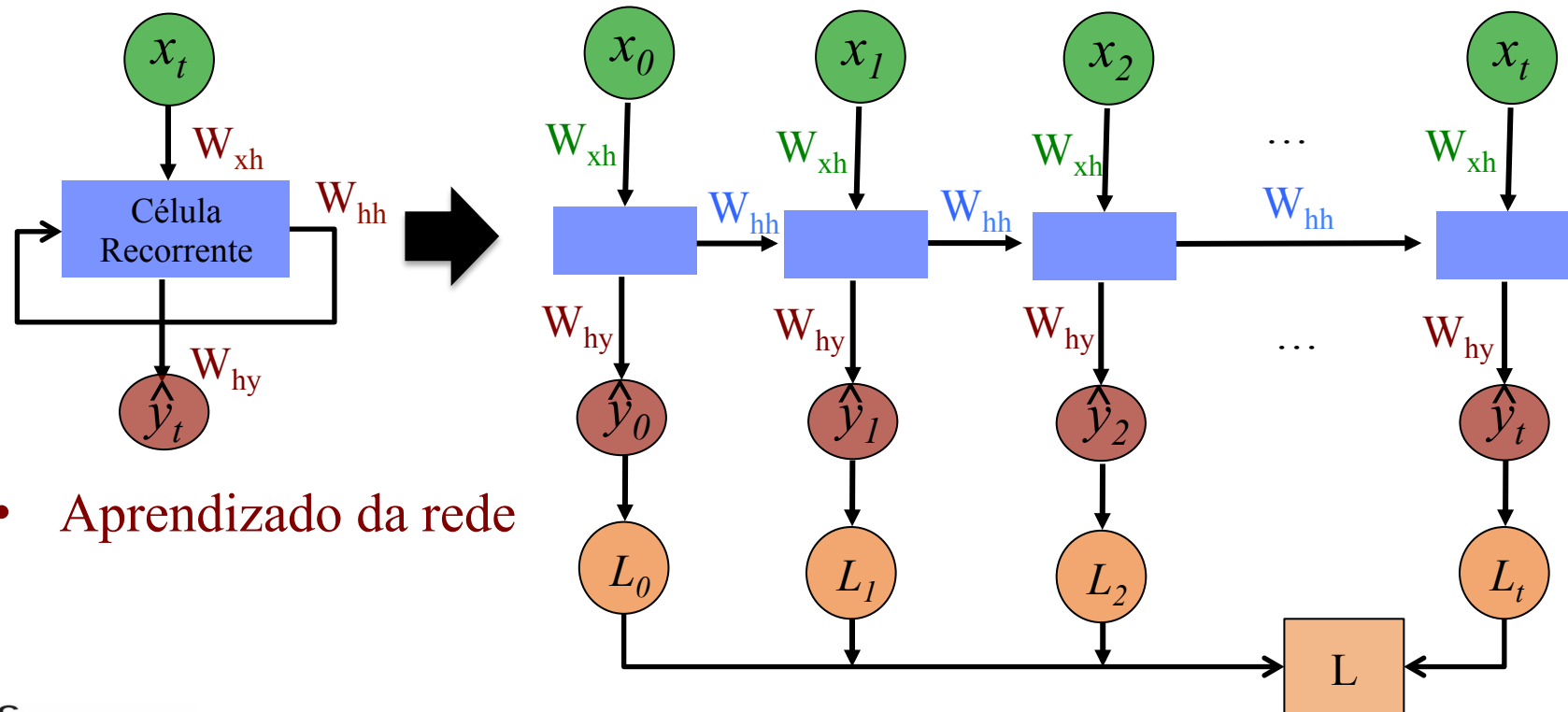
- Treinamento da rede



$$h_t = \tanh(W_{hh} h_{t-1} + W_{xh} x_t)$$

$$\hat{y}_t = W_{hy} h_t$$

Redes Neurais Recorrentes (RNN)



- Aprendizado da rede

Previsão da próxima palavra de uma sequência

Estou cansada de ficar em casa

Entrada

Saída

- É preciso gerar uma representação numérica para as palavras
 1. Conhecer o vocabulário da base de dados sendo utilizada
 2. Mapear cada palavra do vocabulário em um índice
 3. Gerar um embedding – transformar esse índice em um vetor de tamanho fixo

Previsão da próxima Palavra de uma Sequência

Estou cansada de ficar em casa

cansada	cansada	1	“cansada” = [0 1 0 0 0 0]
ficar	ficar	2	
...	
de	de	n	
Vocabulário	Índice	One-hot encoding	

Requisitos de Modelos de Sequência

1. Lidar com sequências de tamanho variável
2. Tratar dependências de longo prazo
3. Manter informações da ordem dos termos da sequência
4. Compartilhar parâmetros ao longo da sequência

Requisitos de Modelos de Sequência

1. Lidar com sequências de tamanho variável

Estou cansada.

Estou cansada de ficar em casa.

Eles fizeram uma trilha no fim de semana e por isso estão cansados.

Requisitos de Modelos de Sequência

2. Tratar dependências de longo prazo

Júlia levou Clara a Disney nas últimas férias e **ela** gostou muito.

Requisitos de Modelos de Sequência

3. Manter informações da ordem dos termos da sequência

The food was good, not bad at all.

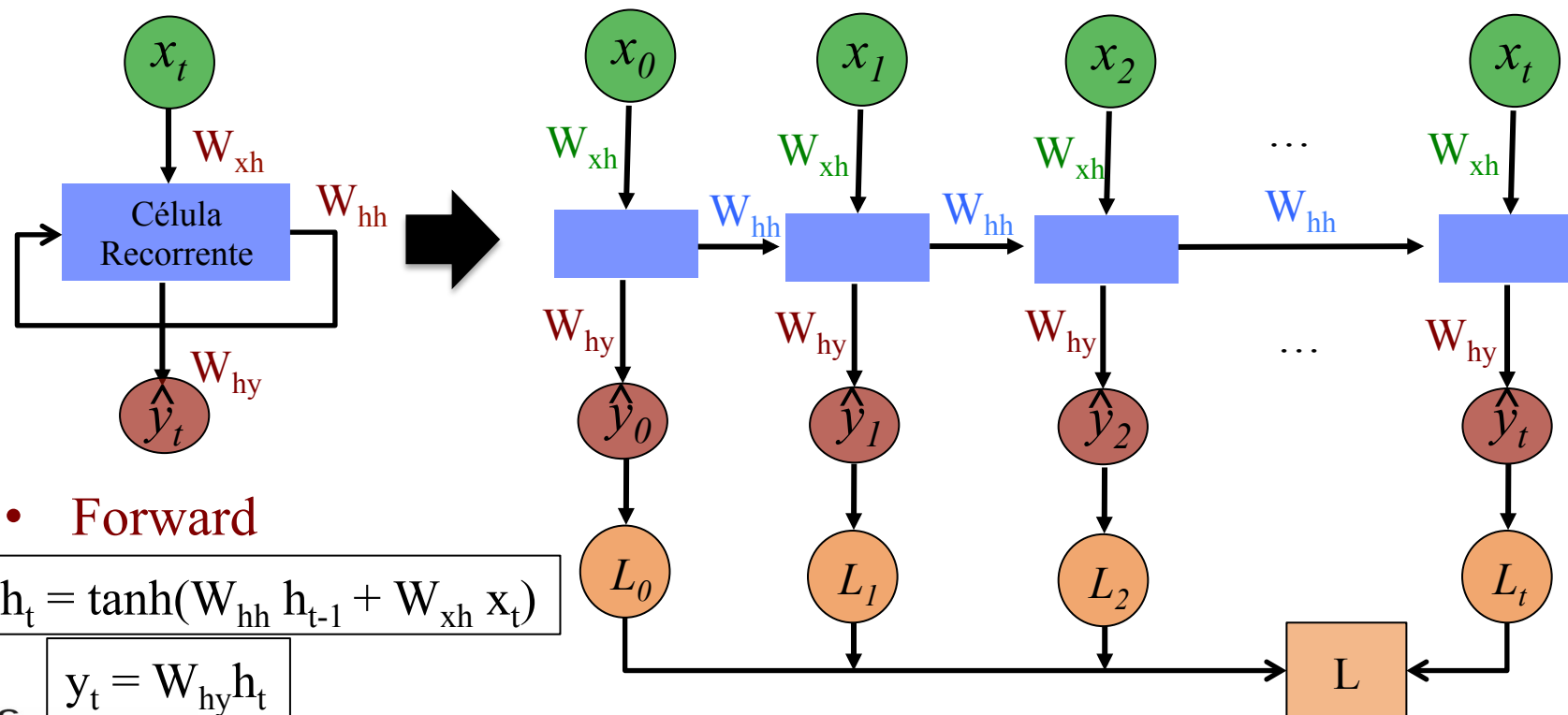
The food was bad, not good at all.

Requisitos de Modelos de Sequência

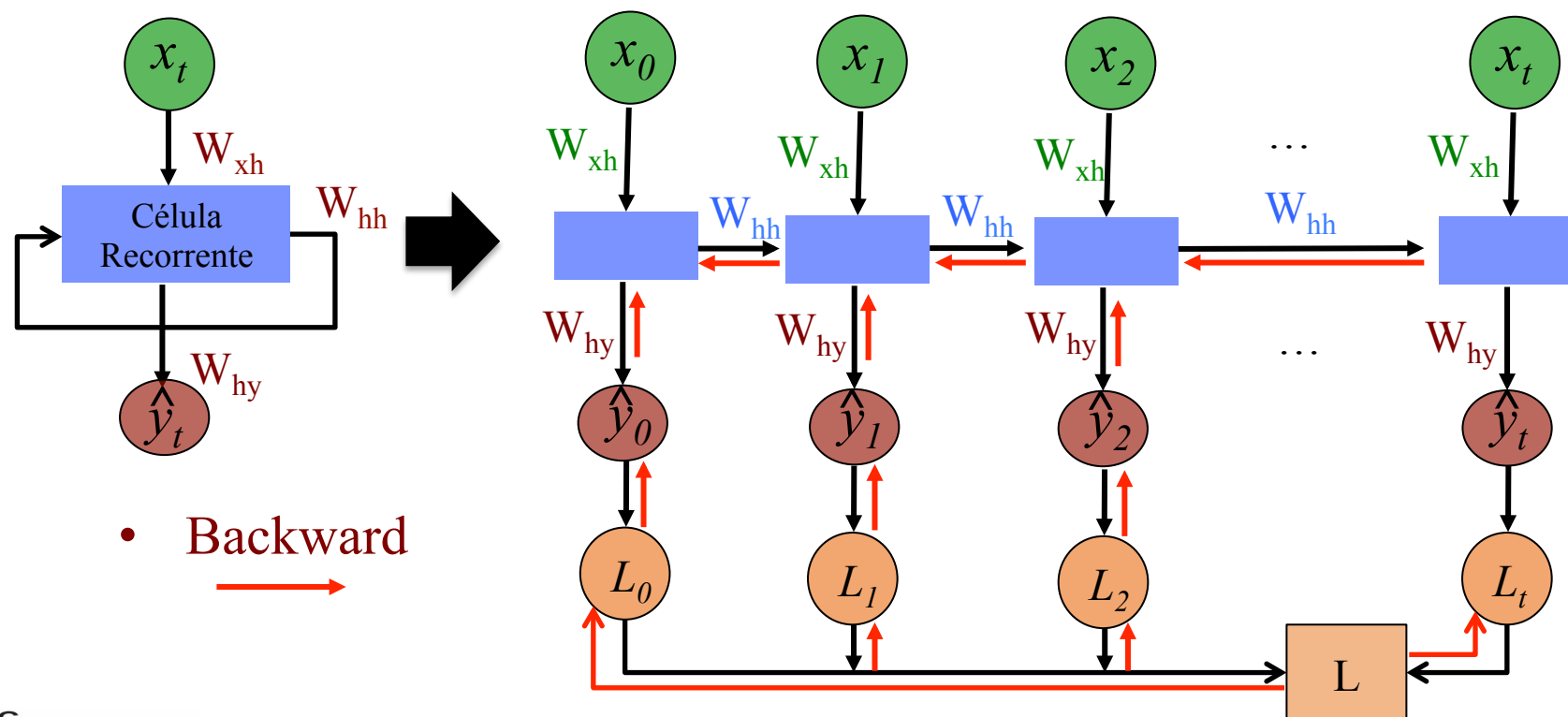
1. Lidar com sequências de tamanho variável
2. Tratar dependências de longo prazo
3. Manter informações da ordem dos termos da sequência
4. Compartilhar parâmetros ao longo da sequência

Redes Recorrentes cumprem todos
esses requisitos!

RNNs: Backpropagation through time

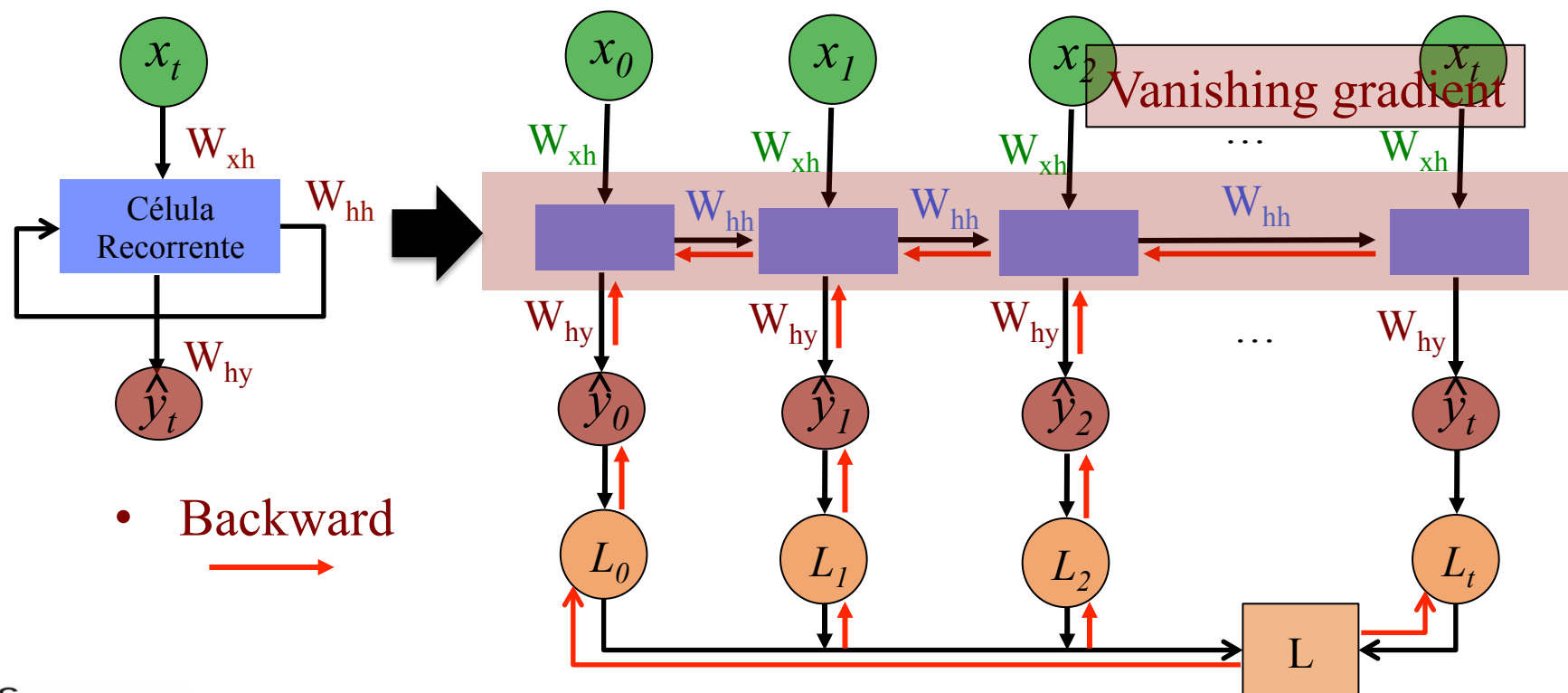


RNNs: Backpropagation through time



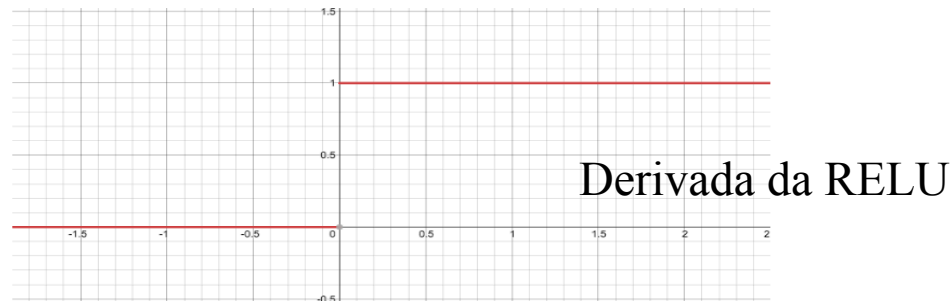
• Backward

RNNs: Backpropagation through time



Soluções para problema do *vanishing gradient*

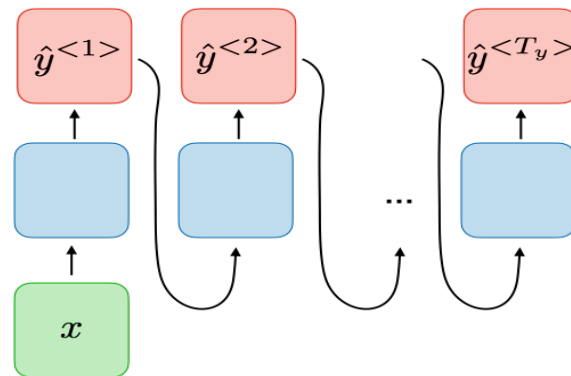
- Uso da RELU



- Inicialização inteligente dos parâmetros
- Uso de células recorrentes mais poderosas, que utilizam o conceito de portas (*gates*)
 - LSTM, GRU

Redes Neurais Recorrentes (RNN)

- Back propagation through time (BPTT) é utilizado quando temos redes com ciclos na camada escondida
- Redes com ciclos da camada de saída para a escondida usam outros tipos de algoritmos (ex: *teacher forcing*)



Um para muitos
Criação de legendas de imagens

Referências

- <http://www.deeplearningbook.org/>
- Aula do curso do MIT 6.S191: “Introduction to Deep Learning”