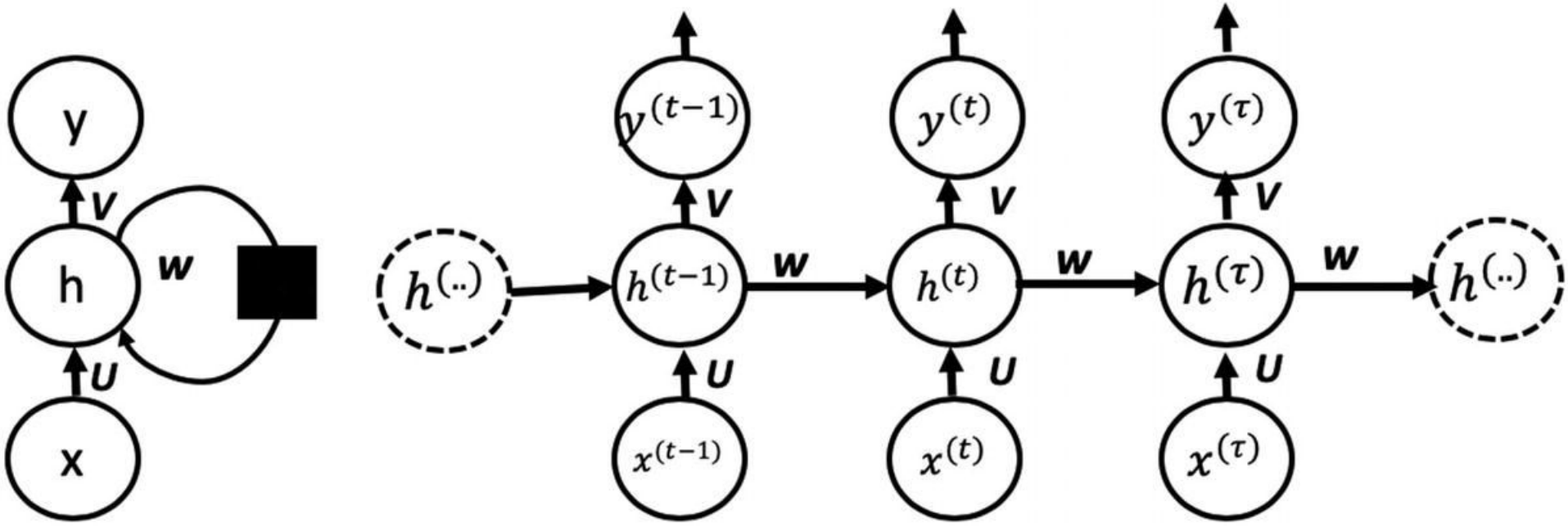
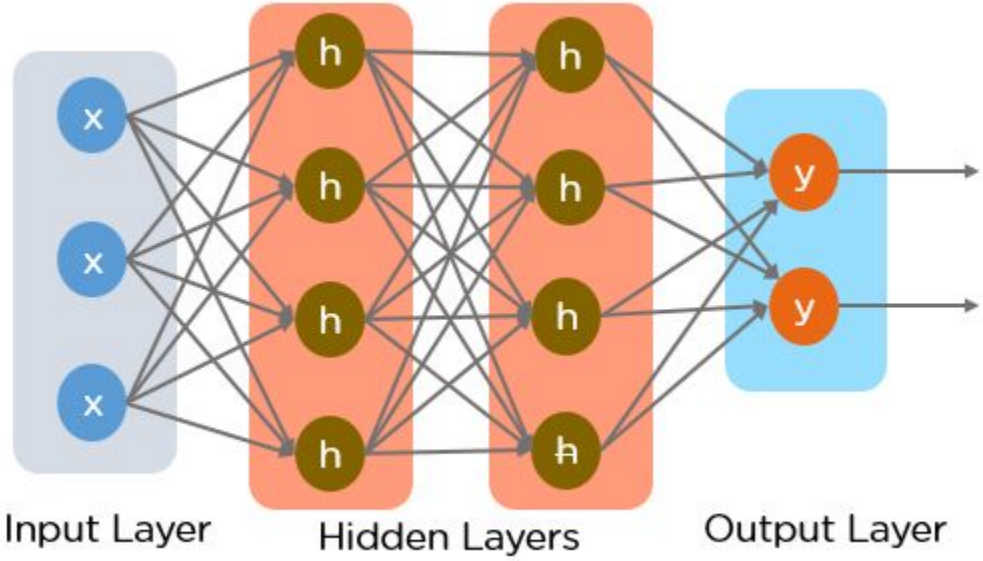
Recurrent Neural Networks

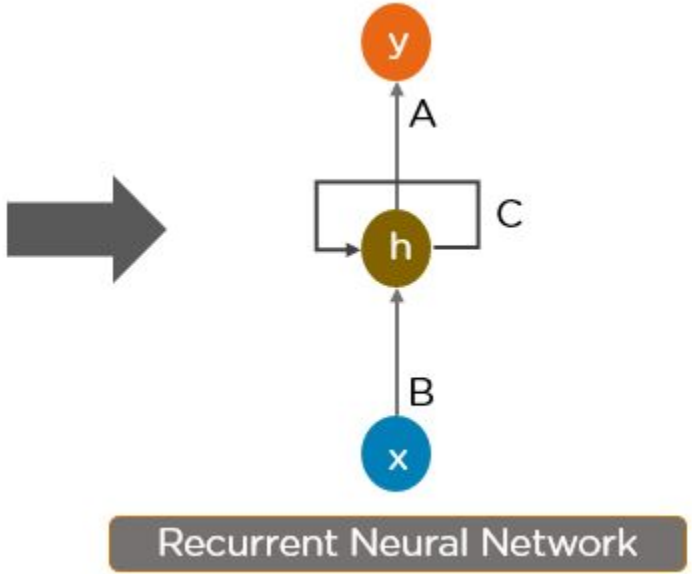
Recurrent Neural Networks

Recurrent neural networks (RNNs) are essentially neural networks that employ recurrence, which is using information from a previous forward pass over the neural network.



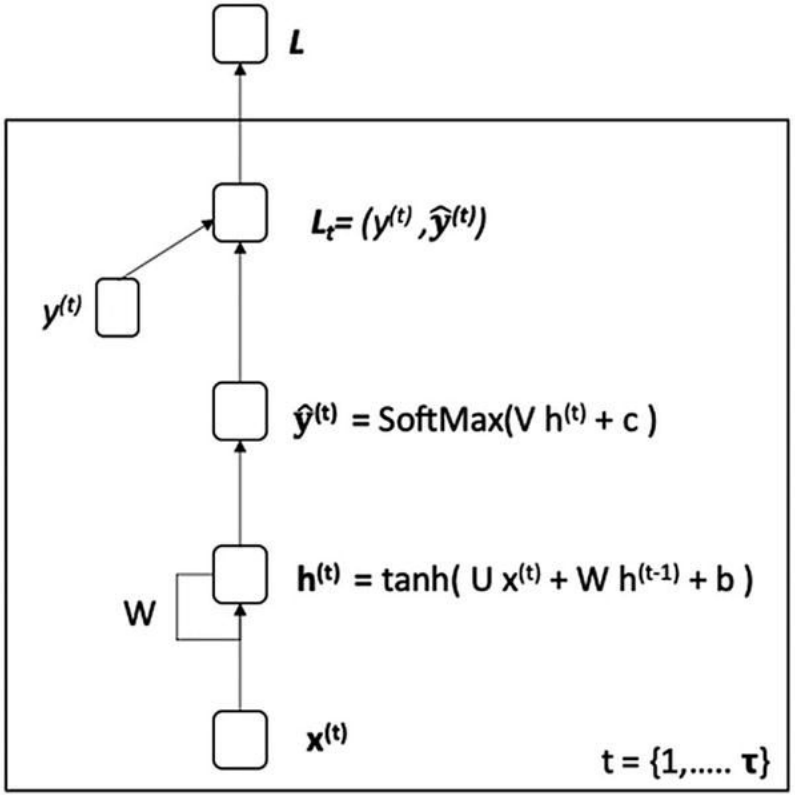
Recurrent Neural Networks

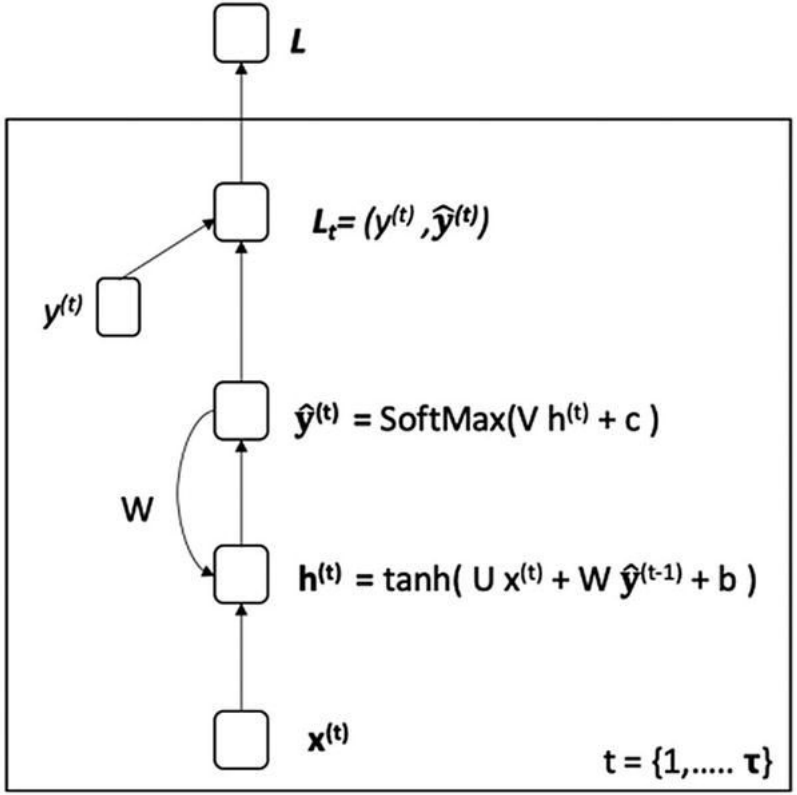




RNN -Redes Neurais Recorrentes

Recurrent Neural Networks

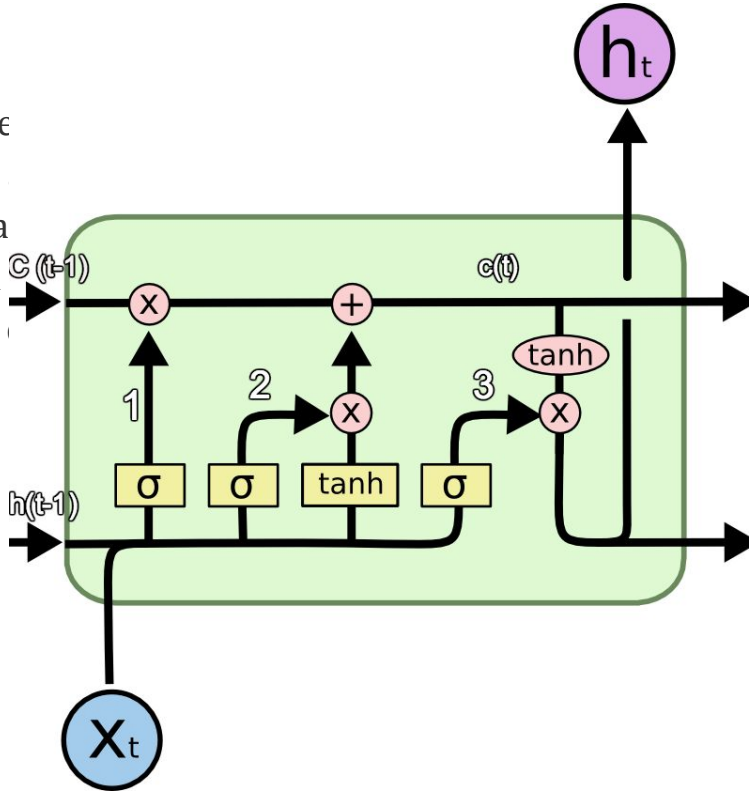




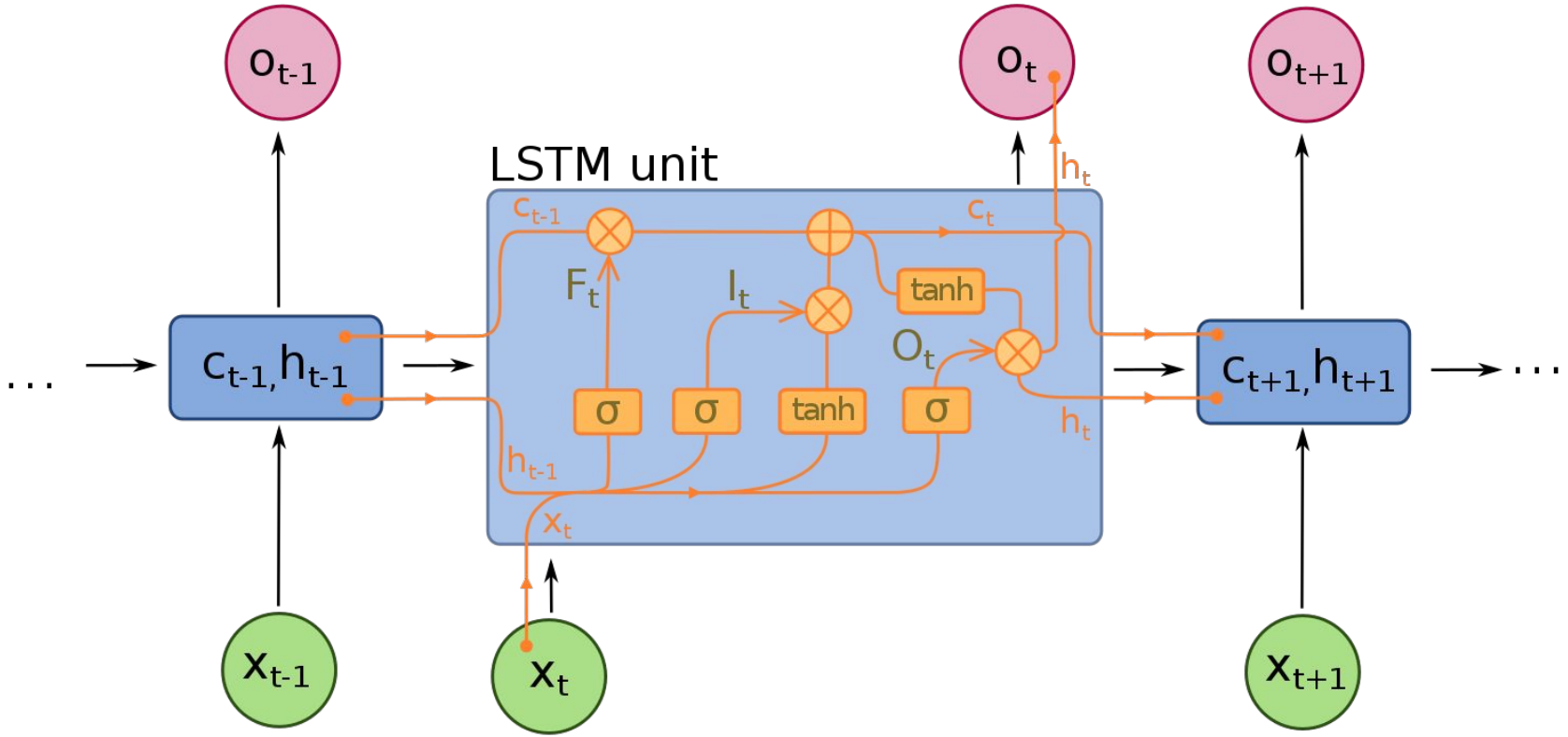
Recurrent Neural Networks

LSTM - Long-Short Term Memory

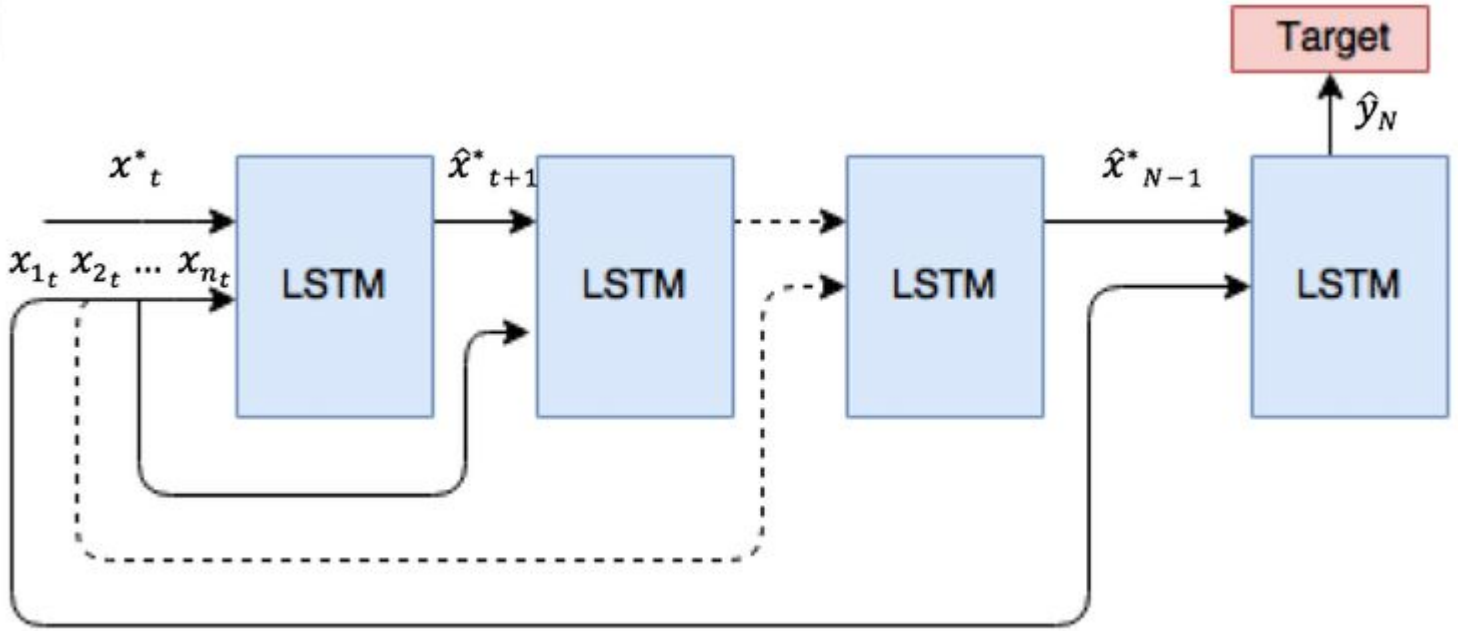
The core concepts within LSTM networks are the cell state and gate
gates and the cell state include several operations, such as sigmoid
multiplication and addition, and vector concatenation. These opera
the network to forget or propagate important information through
information throughout the network and thus help in passing long
when needed.



Recurrent Neural Networks



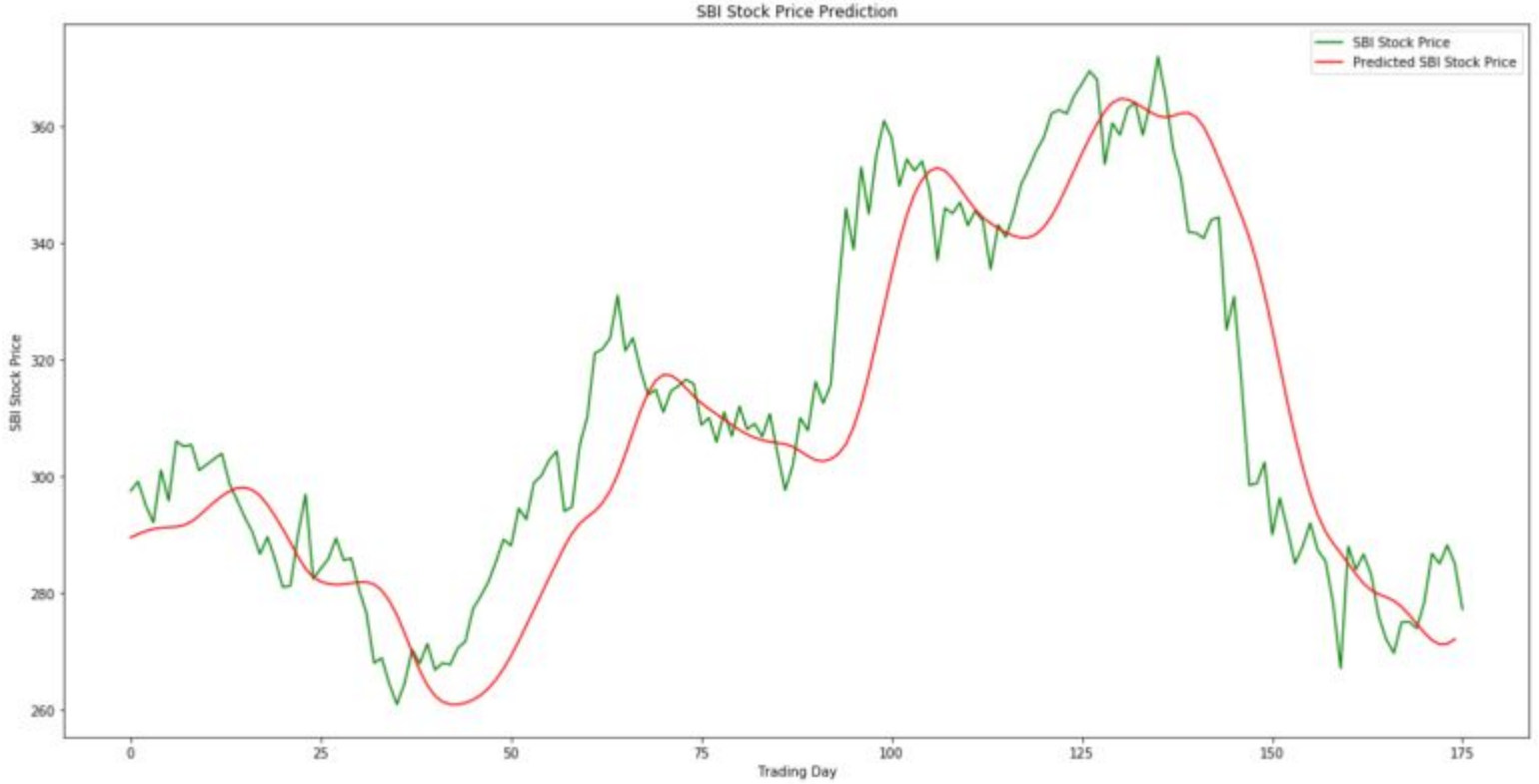
Recurrent Neural Networks



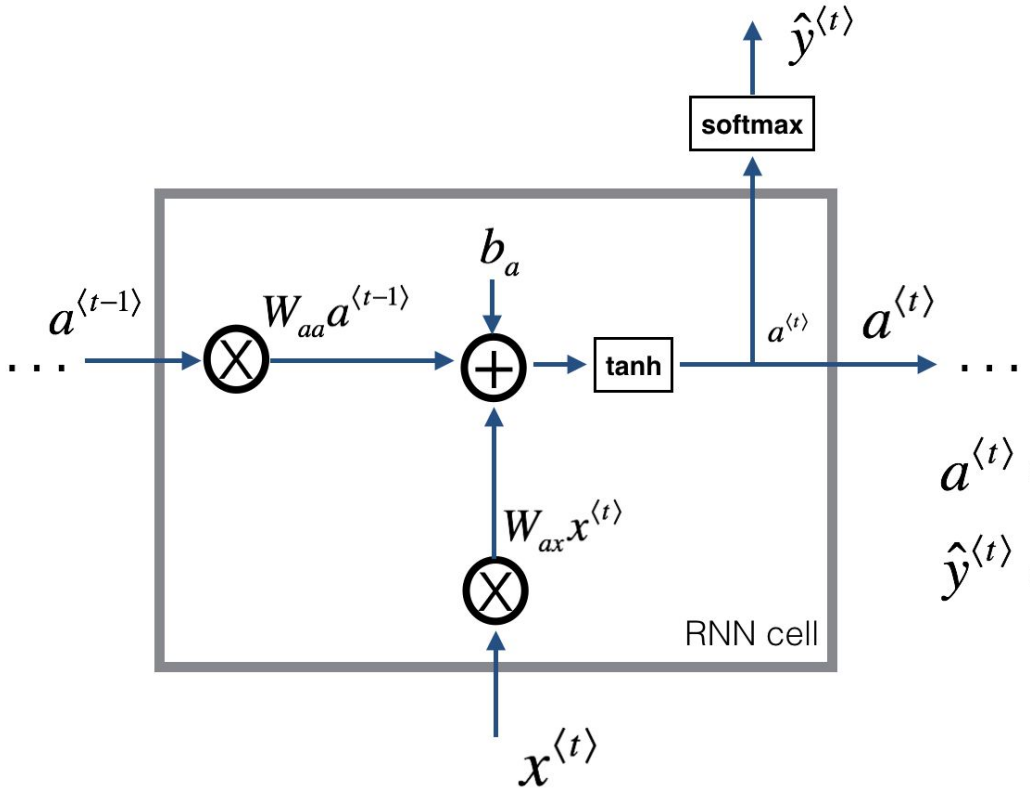
Recurrent Neural Networks

```
[10 20 30] 40
[20 30 40] 50
[30 40 50] 60
[40 50 60] 70
[50 60 70] 80
[60 70 80] 90
```

Recurrent Neural Networks



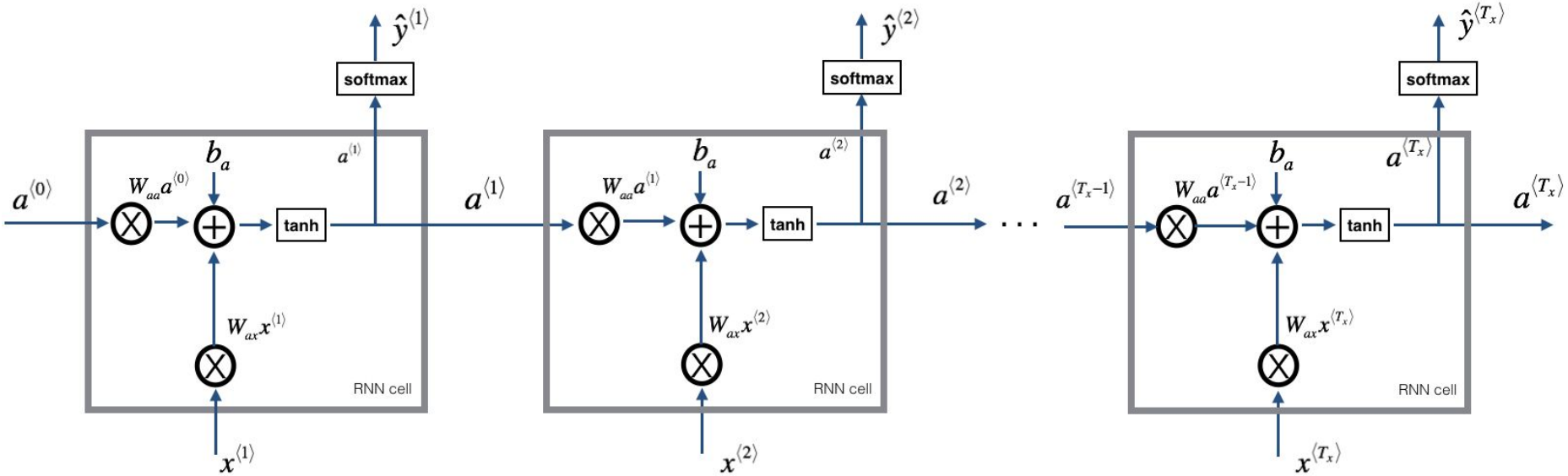
RNN



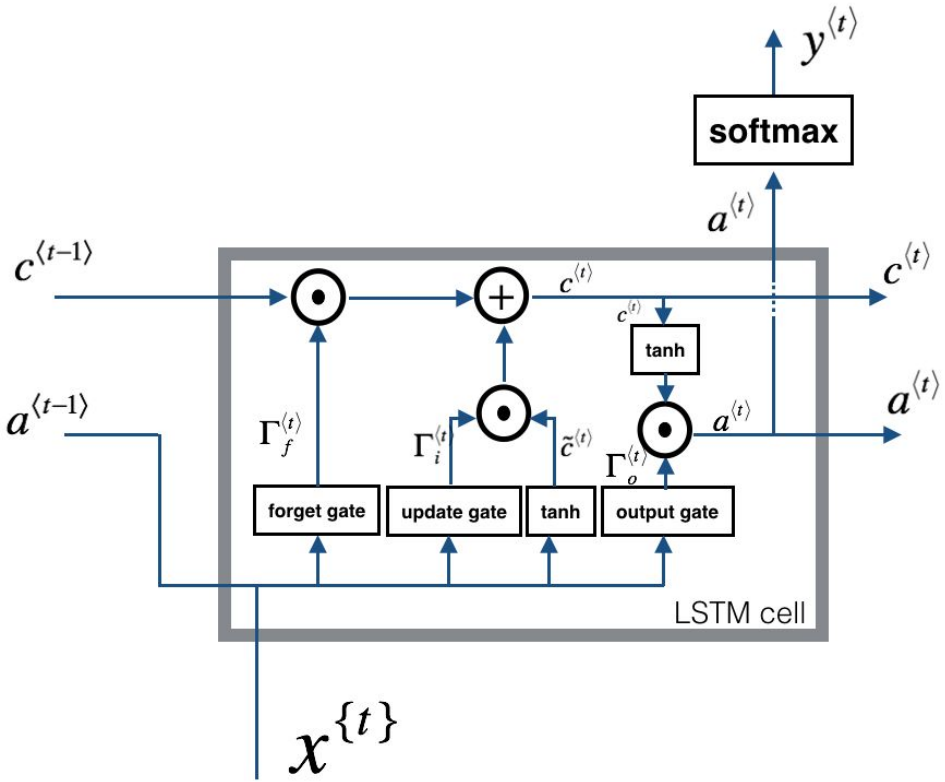
$$a^{\langle t \rangle} = \tanh(W_{ax}x^{\langle t \rangle} + W_{aa}a^{\langle t-1 \rangle} + b_a)$$

$$\hat{y}^{\langle t \rangle} = soft\max(W_{ya}a^{\langle t \rangle} + b_y)$$

RNN


$\hat{y}^{\langle 1 \rangle}$
softmax
$a^{\langle 0 \rangle}$
$W_{aa}a^{\langle 0 \rangle}$
b_a
$a^{\langle 1 \rangle}$
tanh
$W_{ax}x^{\langle 1 \rangle}$
RNN cell
$x^{\langle 1 \rangle}$
$a^{\langle 1 \rangle}$
$\hat{y}^{\langle 2 \rangle}$
softmax
$W_{aa}a^{\langle 1 \rangle}$
b_a
$a^{\langle 2 \rangle}$
tanh
$W_{ax}x^{\langle 2 \rangle}$
RNN cell
$x^{\langle 2 \rangle}$
$a^{\langle 2 \rangle}$
$a^{\langle T_x-1 \rangle}$
$\hat{y}^{\langle T_x \rangle}$
softmax
$W_{aa}a^{\langle T_x-1 \rangle}$
b_a
$a^{\langle T_x \rangle}$
tanh
$W_{ax}x^{\langle T_x \rangle}$
RNN cell
$x^{\langle T_x \rangle}$
$a^{\langle T_x \rangle}$

LSTM



$$\Gamma_f^{\langle t \rangle} = \sigma(W_f[a^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_f)$$

$$\Gamma_u^{\langle t \rangle} = \sigma(W_u[a^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_u)$$

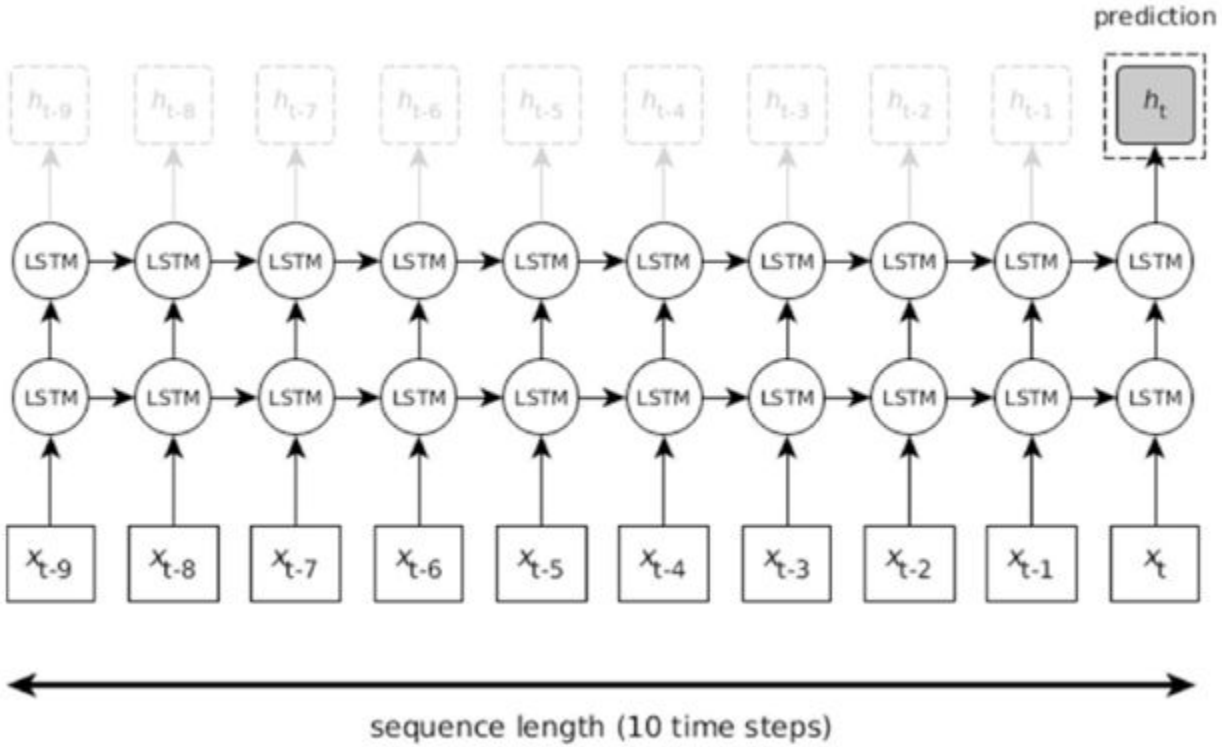
$$\tilde{c}^{\{t\}} = \tanh(W_C[a^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_C)$$

$$c^{\langle t \rangle} = \Gamma_f^{\langle t \rangle} \circ c^{\langle t-1 \rangle} + \Gamma_u^{\langle t \rangle} \circ \tilde{c}^{\langle t \rangle}$$

$$\Gamma_o^{\langle t \rangle} = \sigma(W_o[a^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_o)$$

$$a^{\langle t \rangle} = \Gamma_o^{\langle t \rangle} \circ \tanh(c^{\langle t \rangle})$$

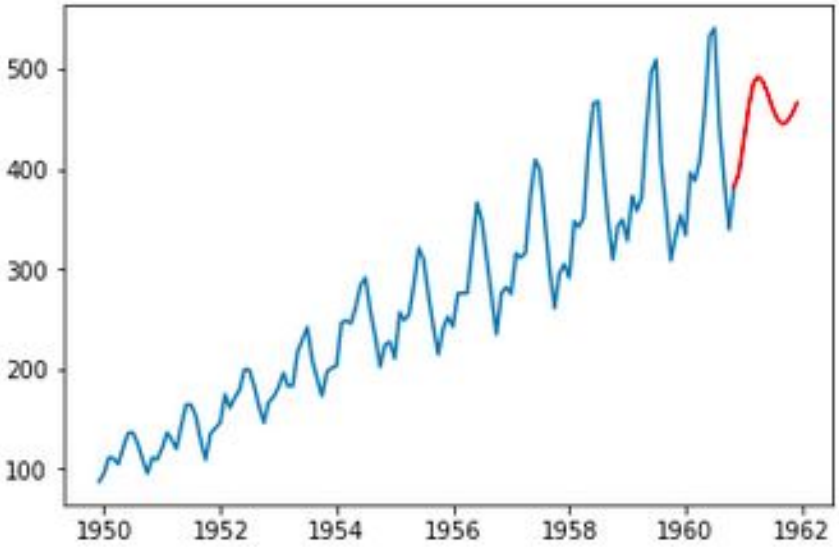
LSTM


prediction
h_{t-9}
h_{t-8}
h_{t-7}
h_{t-6}
h_{t-5}
h_{t-4}
h_{t-3}
h_{t-2}
h_{t-1}
h_t
LSTM
x_{t-9}
x_{t-8}
x_{t-7}
x_{t-6}
x_{t-5}
x_{t-4}
x_{t-3}
x_{t-2}
x_{t-1}
x_t
sequence length (10 time steps)

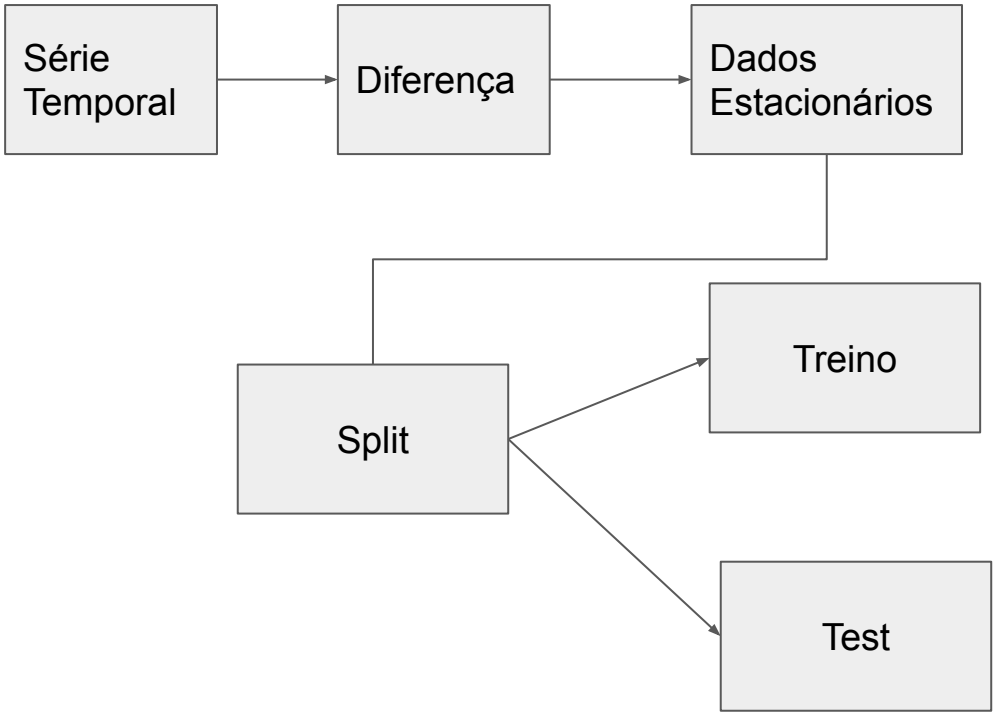
LSTM

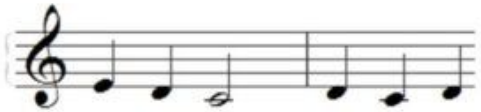
Tendência

Sazonalidade



LSTM


Série Temporal
Diferença
Dados Estacionários
Split
Treino
Test

	x - input		y - output
Speech recognition		⟹	"Fuzzy Wuzzy was a bear. Fuzzy Wuzzy had no hair."
Music generation	∅	⟹	
Sentiment classification	"Decent effort. The plot could have been better."	⟹	
DNA sequence analysis	ACTGTACCCATGTGACTGCCC	⟹	ACTGTACCCATGTGACTGCCC
Machine translation	"El que no arriesga, no gana."	⟹	"If you don't take risks, you cannot win."
Video activity recognition		⟹	Running
Name entity recognition	"Ygritte says Jon Snow knows nothing."	⟹	"Ygritte says Jon Snow knows nothing."