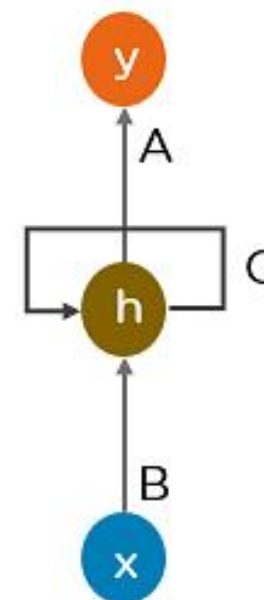
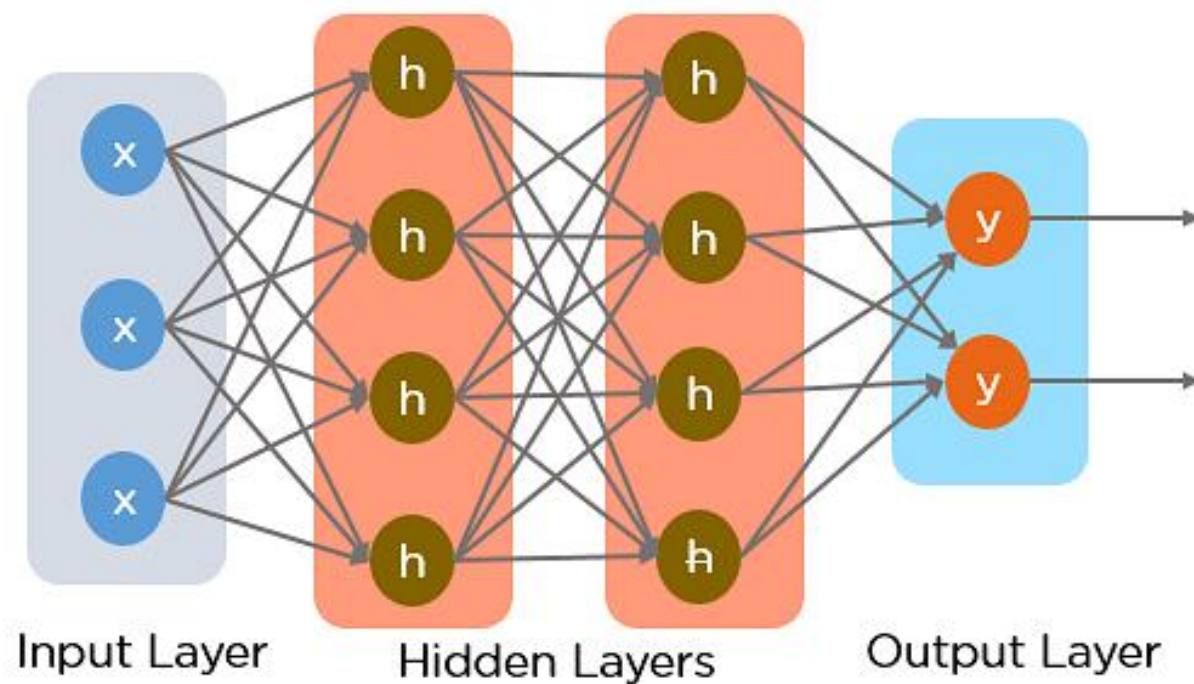


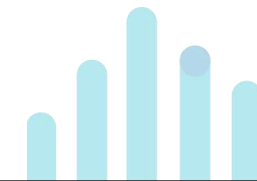
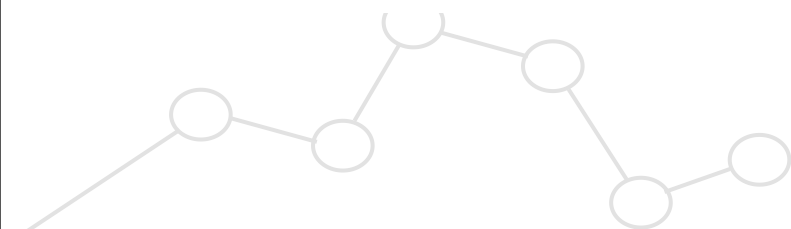
# Recurrent Neural Networks

Dr. Edwyn Javier Aldana Bobadilla

# Recurrent Neural Networks

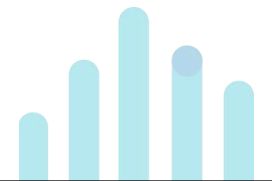
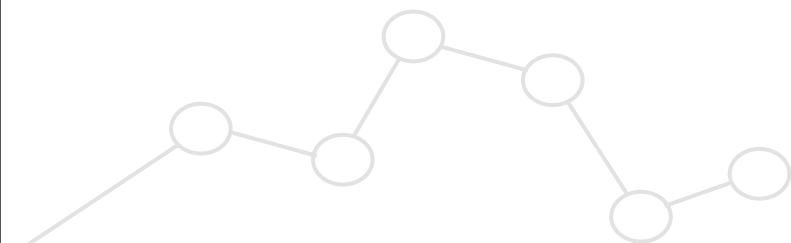


Recurrent Neural Network



# Recurrent Neural Networks

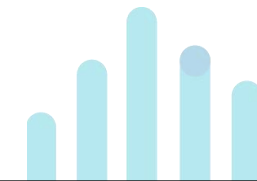
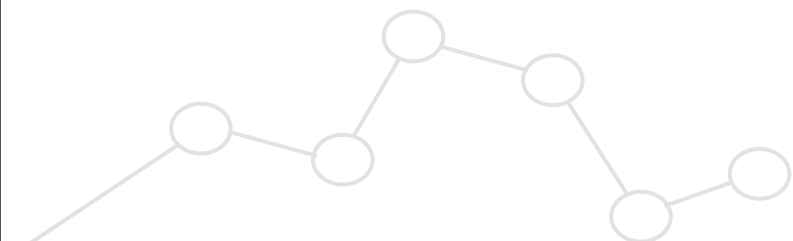
- Recurrent Neural Networks or RNN are a very important variant of neural networks heavily used in **Natural Language Processing**.
- In traditional neural network, an input is processed through a number of layers and an output is produced, with an assumption that two successive **inputs are independent of each other**.
- This assumption is not true in many real-life scenarios: price of a stock at a given time, the  $n^{\text{th}}$  word in a sequence of words.
- The above examples are common application of RNNs.



# Recurrent Neural Networks

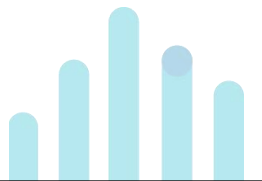
- RNNs are recurrent because they carry out **the same task for each element of a sequence**.
- It is said that a RNN has a **memory** which captures information about what has been calculated previously.

- 

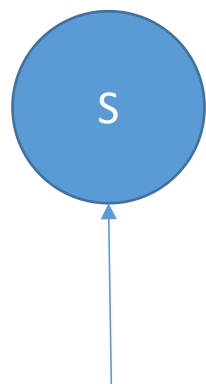


# RNN Architecture

$s = state$



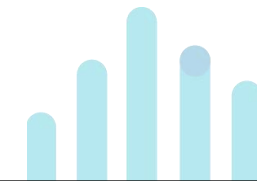
# RNN Architecture



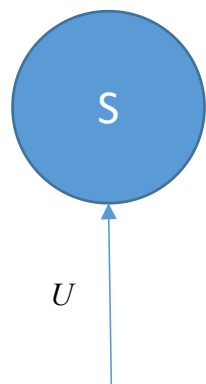
$s = \textit{state}$

$x = \textit{input}$

$x$



# RNN Architecture

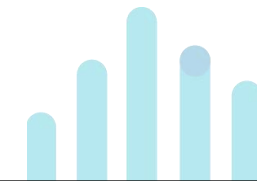


$s = state$

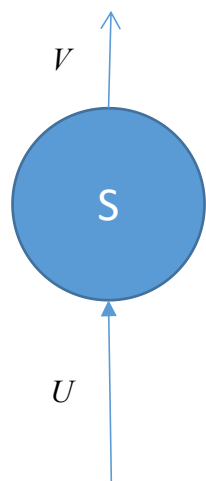
$x = input$

$U = weights\_value$

$x$



# RNN Architecture



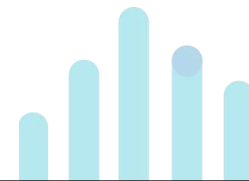
$s = \text{state}$

$x = \text{input}$

$U = \text{weights vector}$

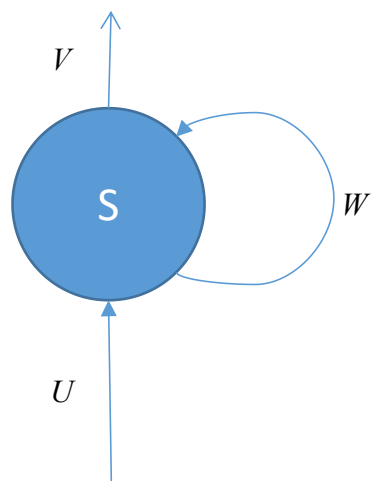
$V = \text{output value}$

$x$





# RNN Architecture



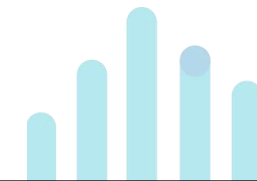
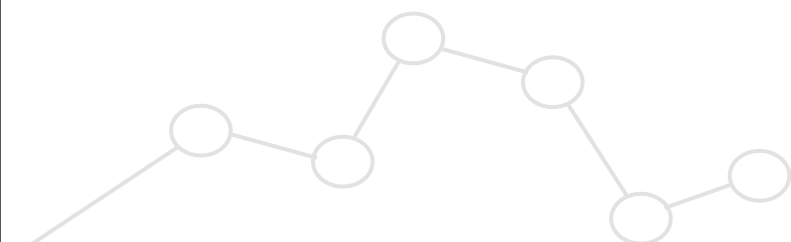
$s = state$

$x = input$

$U = weights\ vectors$

$V, W = output\ values$

$x$



# RNN Architecture

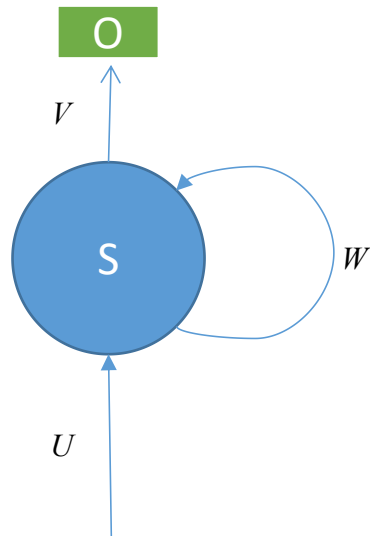
$s = state$

$x = input$

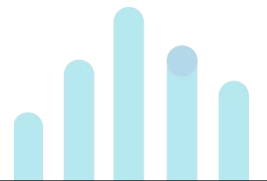
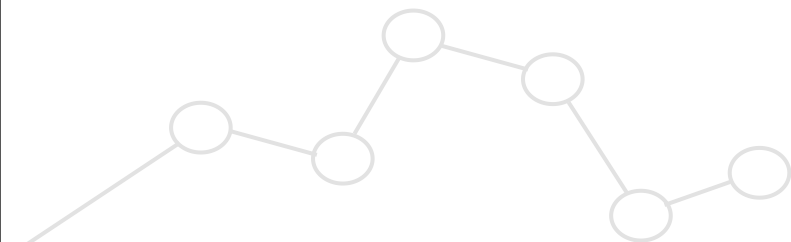
$U = weights vectors$

$V, W = output values$

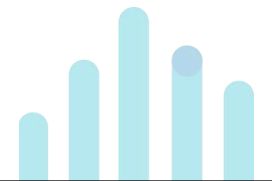
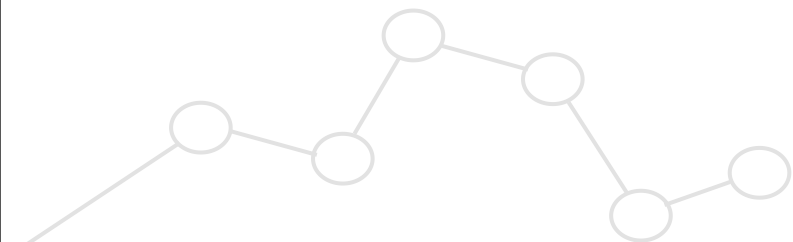
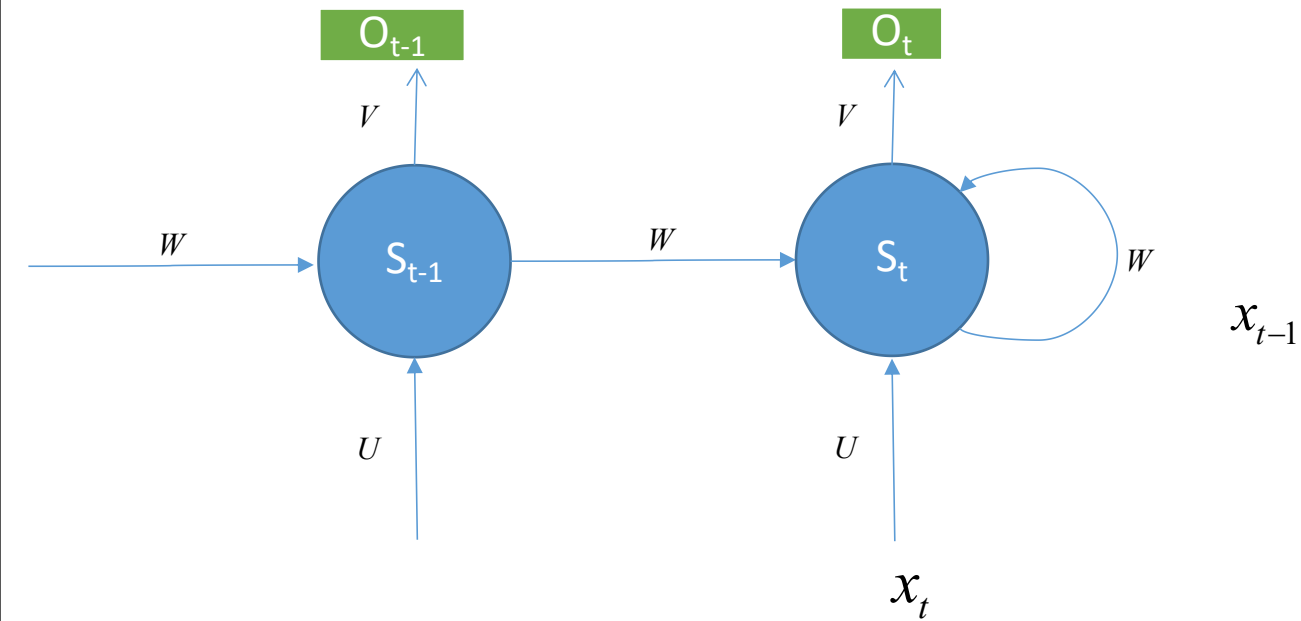
$O = Final Output$



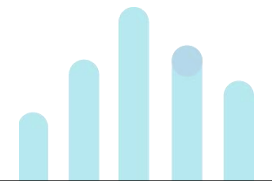
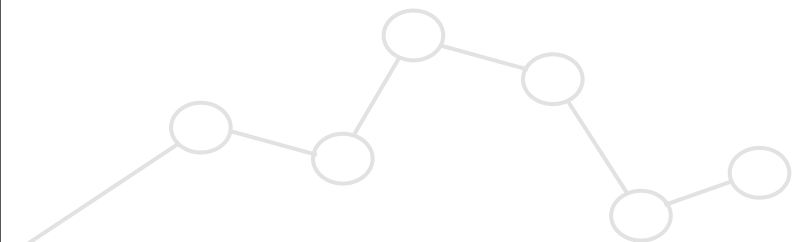
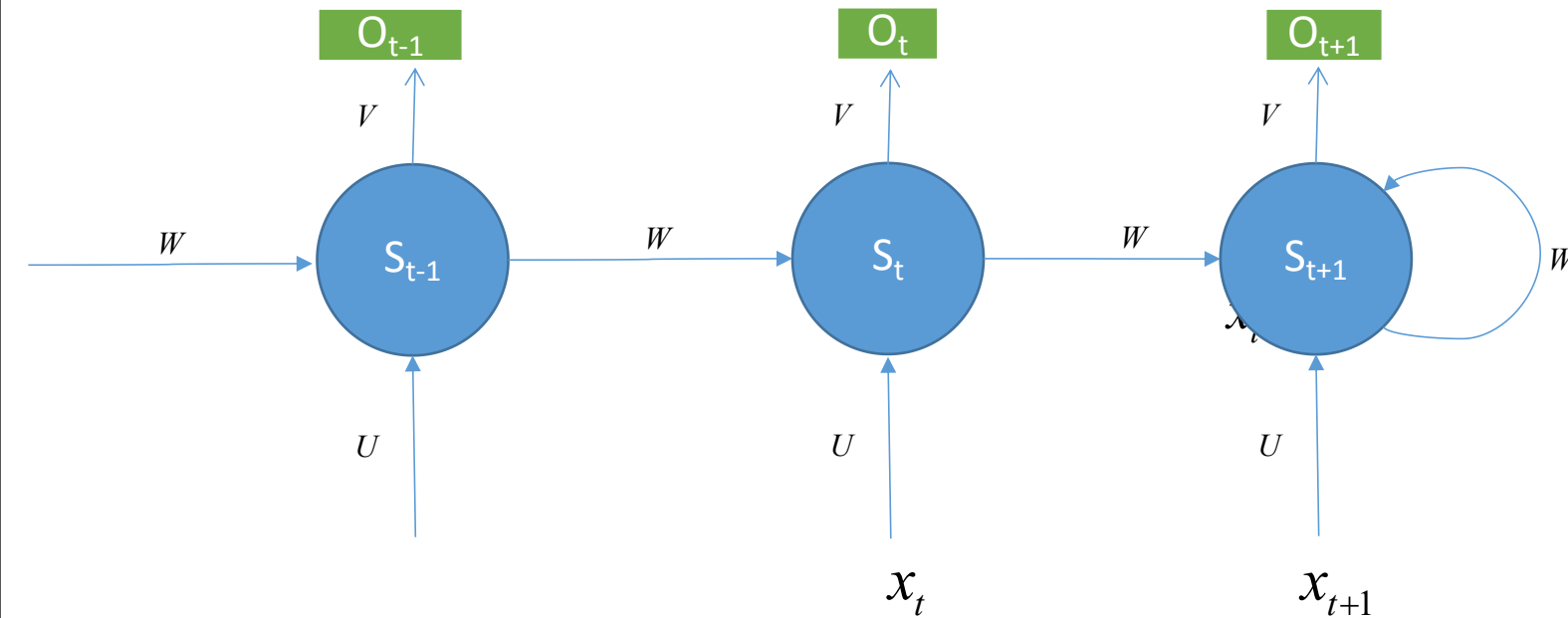
$x$



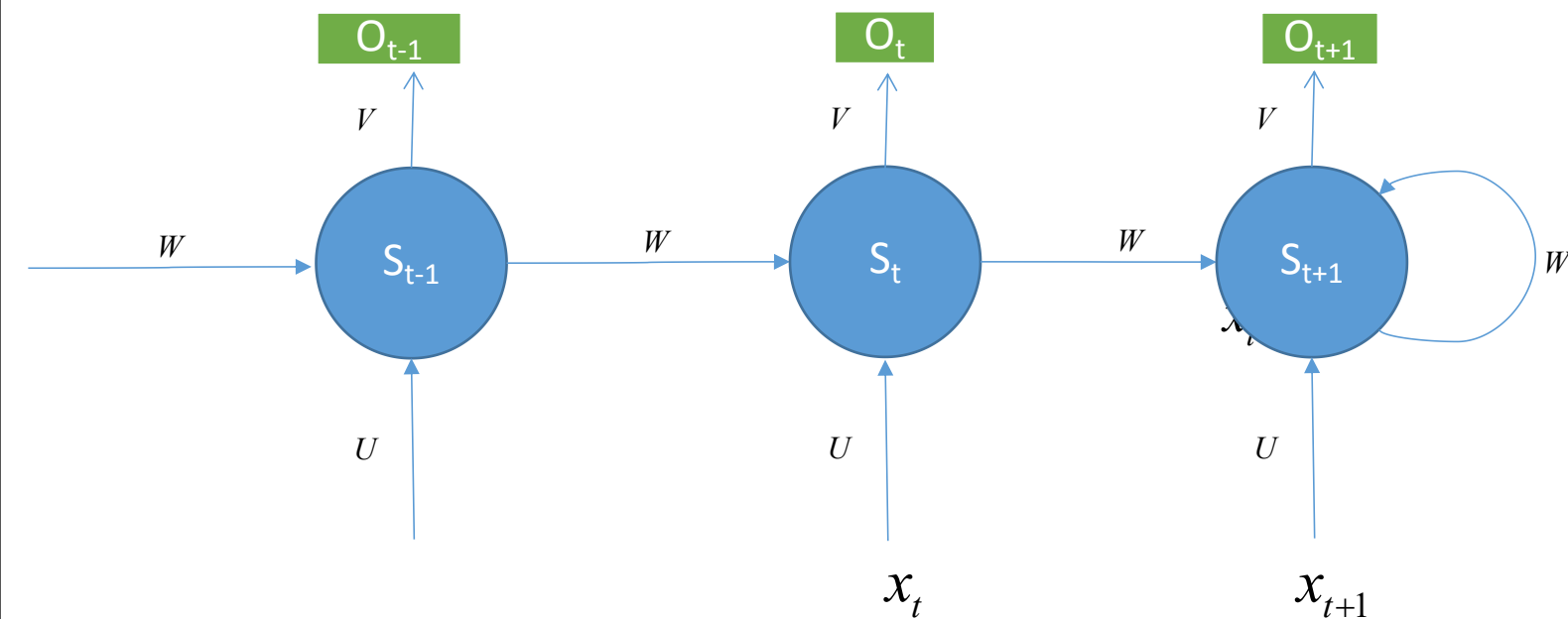
# Unfolding a RNN



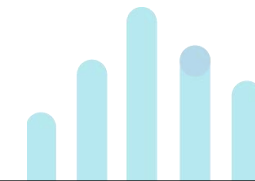
# Unfolding a RNN



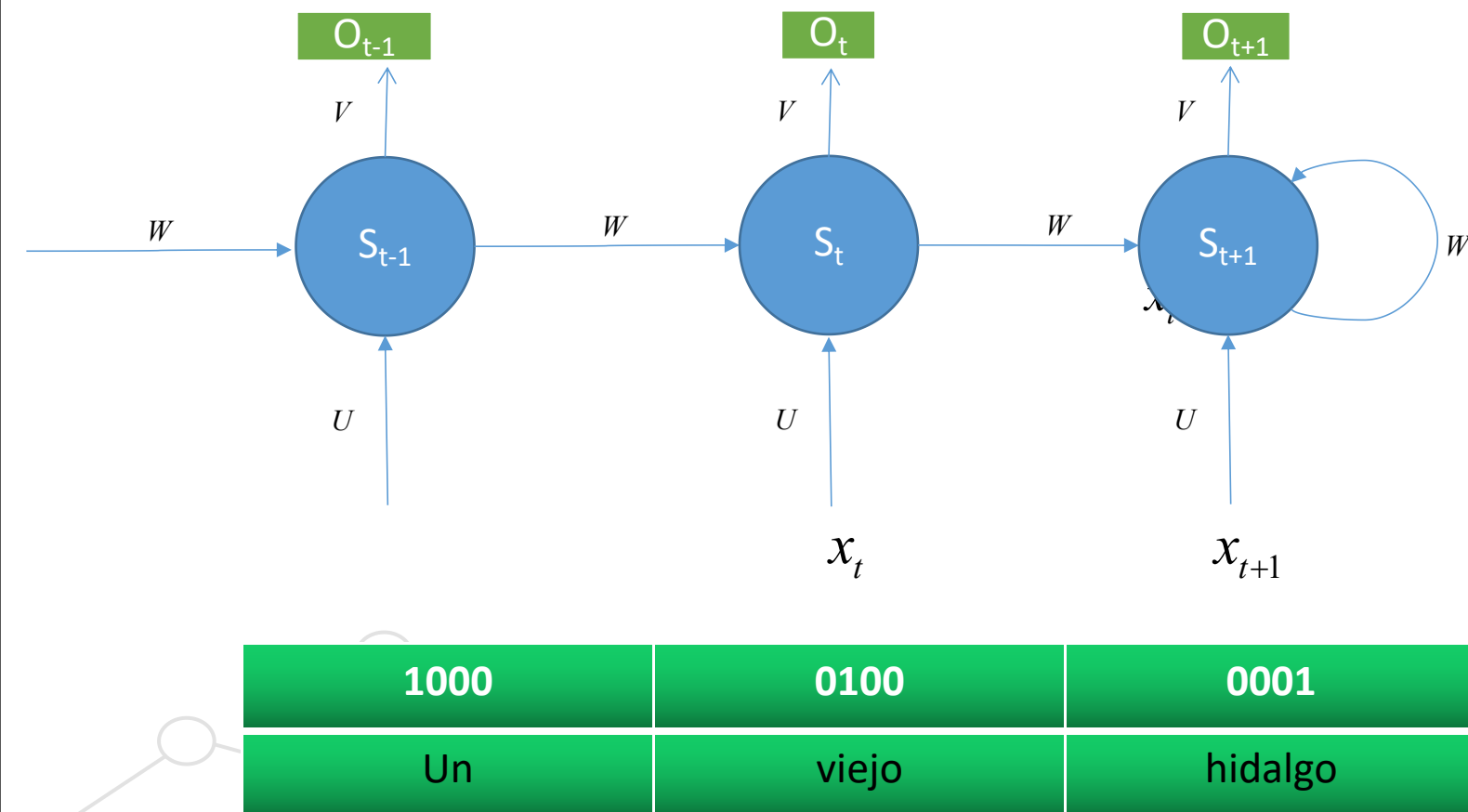
# Example of a RNN



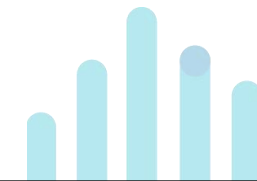
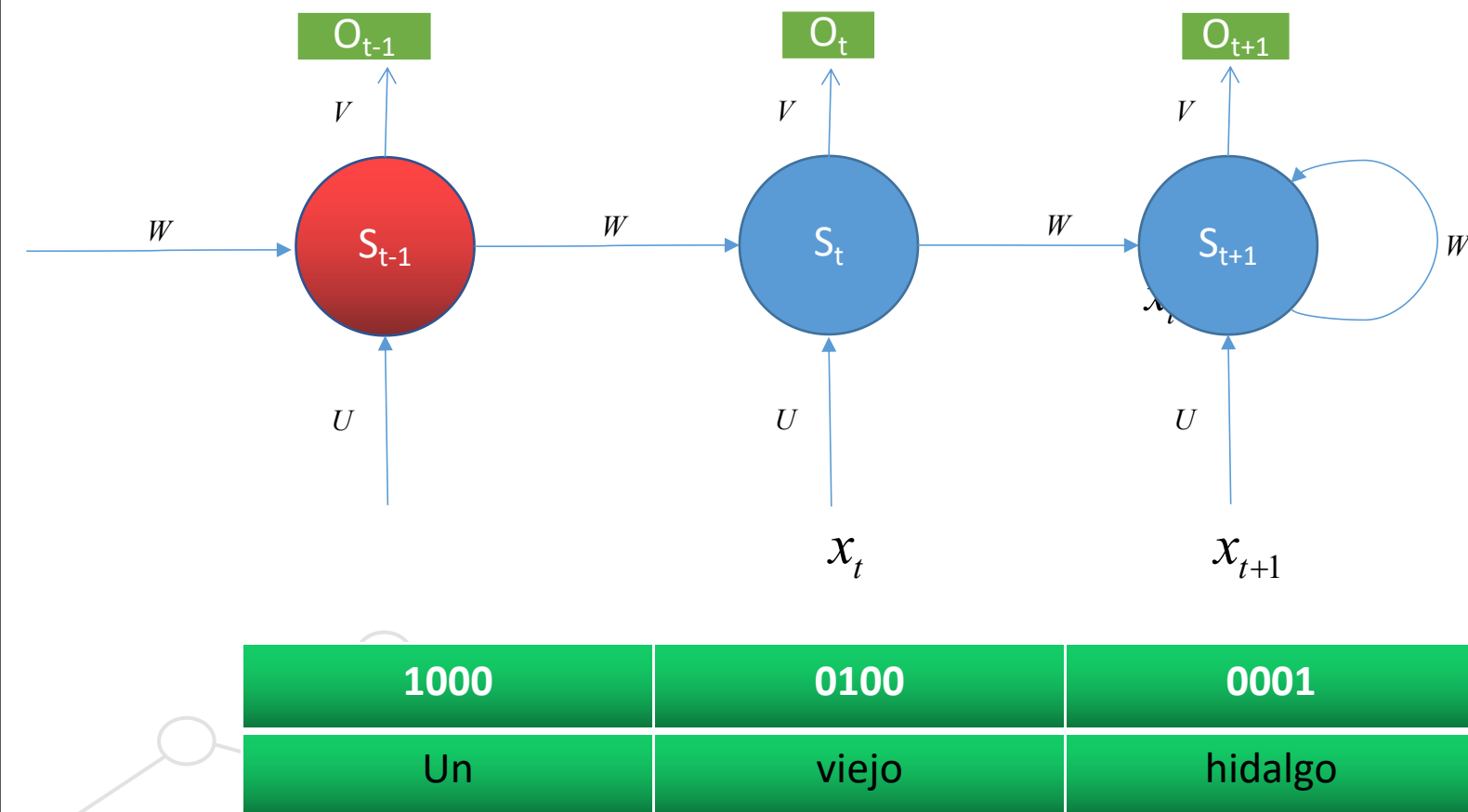
1000	0100	0001
------	------	------



# Example of a RNN



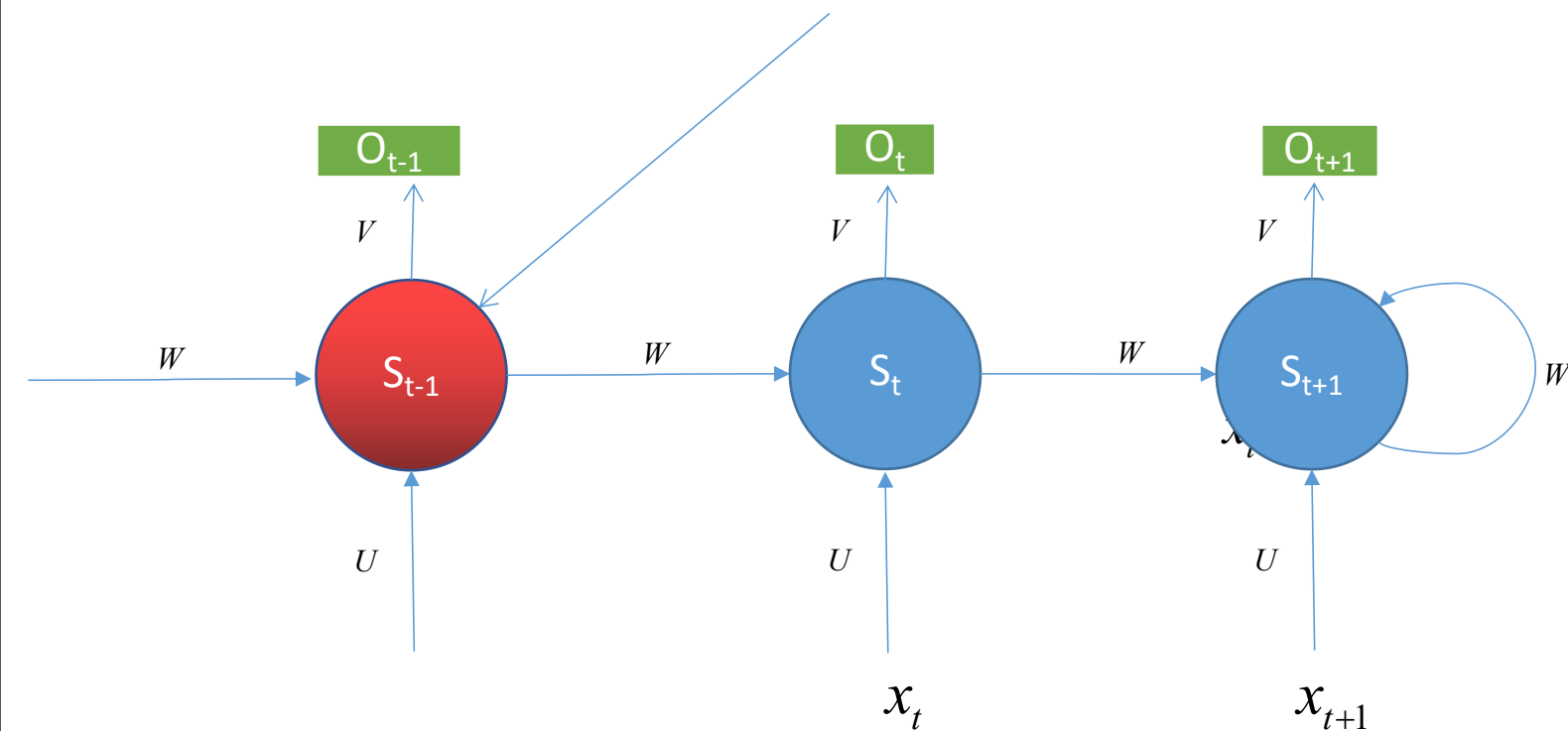
# Example of a RNN



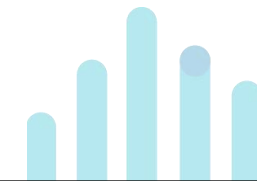
# Example of a RNN

$$s_{t-1} = f(U_{t-1}x_{t-1} + W_{t-1})$$

The function  $f$  is commonly tanh or ReLU



1000	0100	0001
Un	viejo	hidalgo

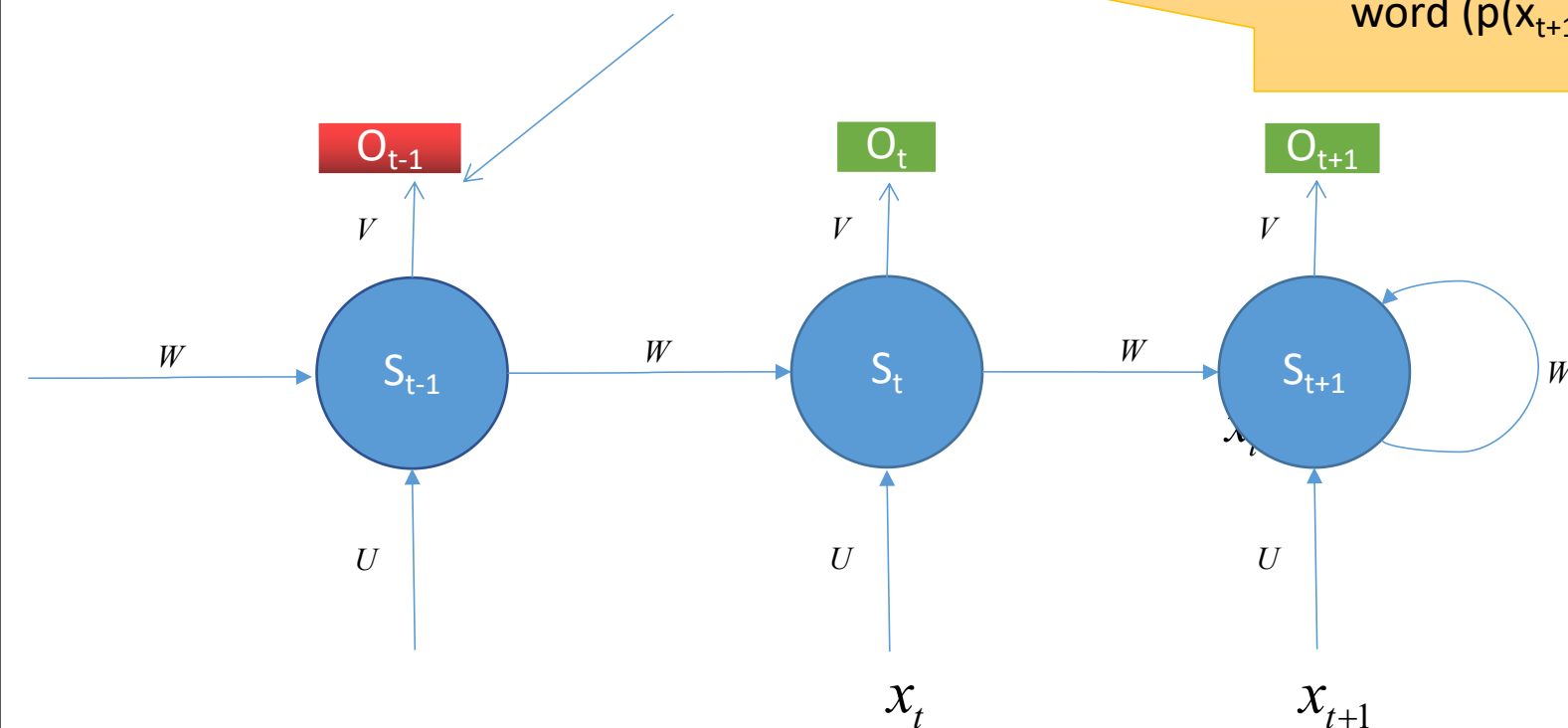




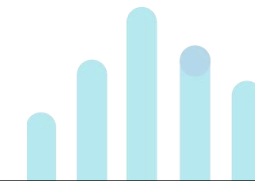
# Example of a RNN

$$o_{t-1} = \text{softmax}(V_{t-1})$$

The softmax function allows us to find the probability of the next word given the actual word ( $p(x_{t+1} | x_t)$ )

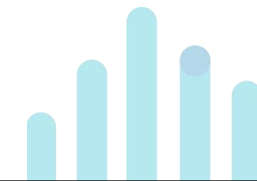
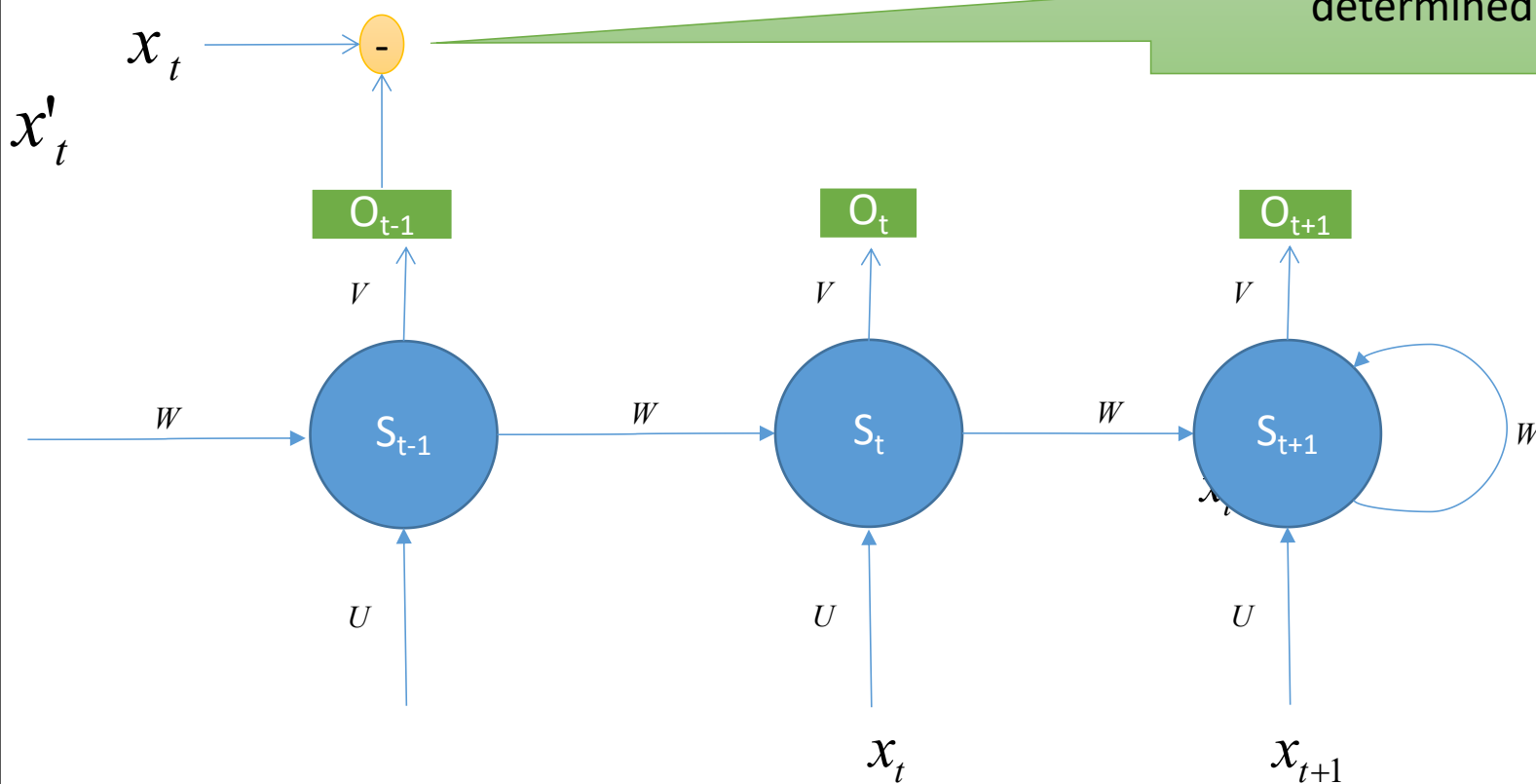


1000	0100	0001
Un	viejo	hidalgo

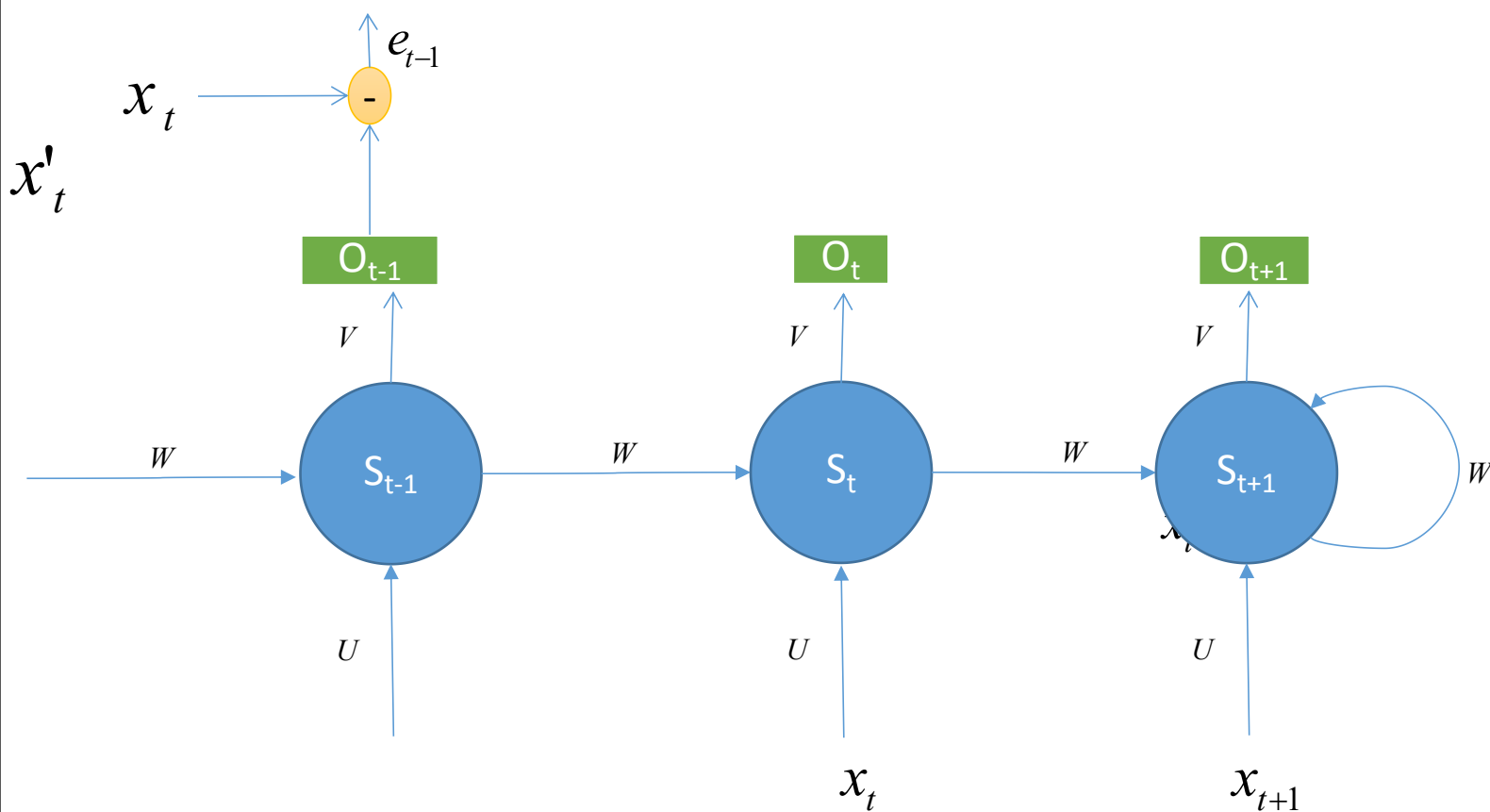


# Training Process

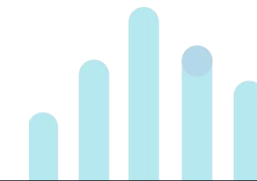
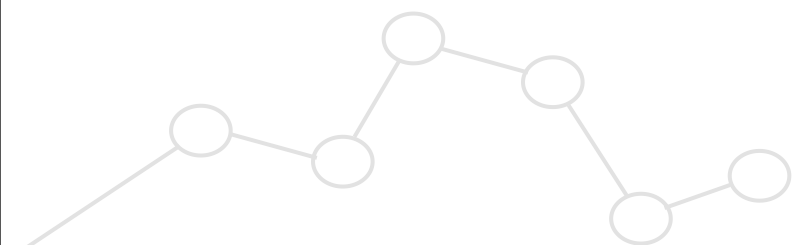
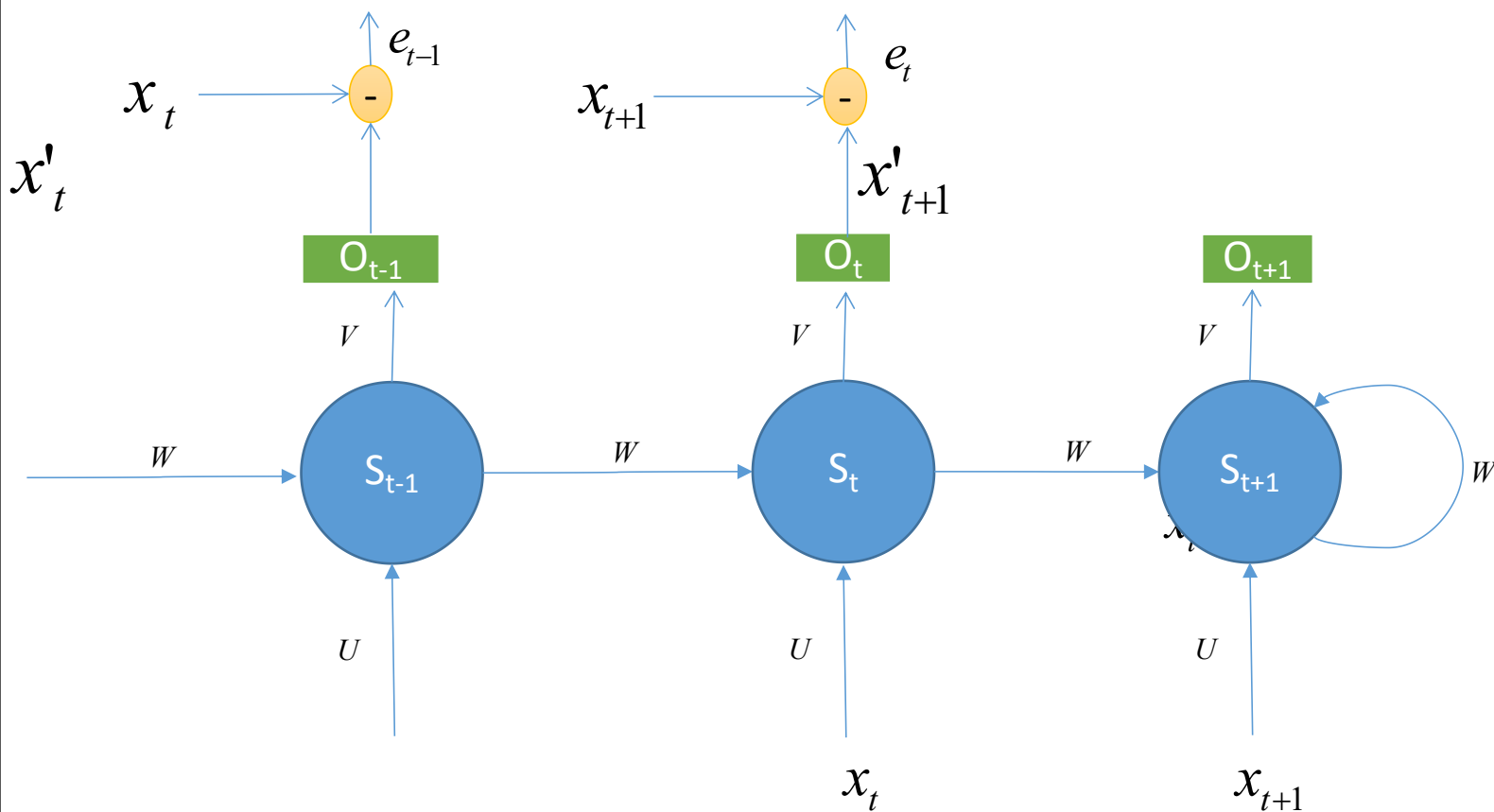
An **error function** can be determined.



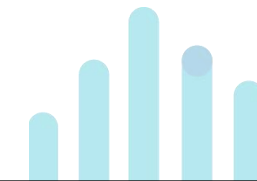
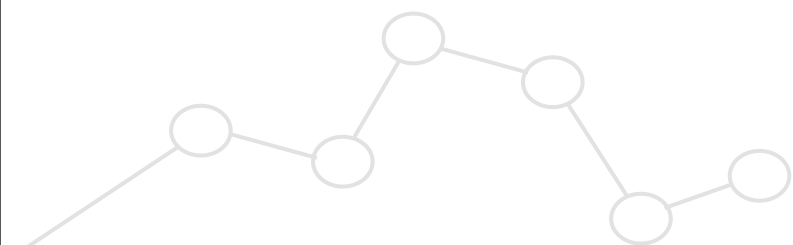
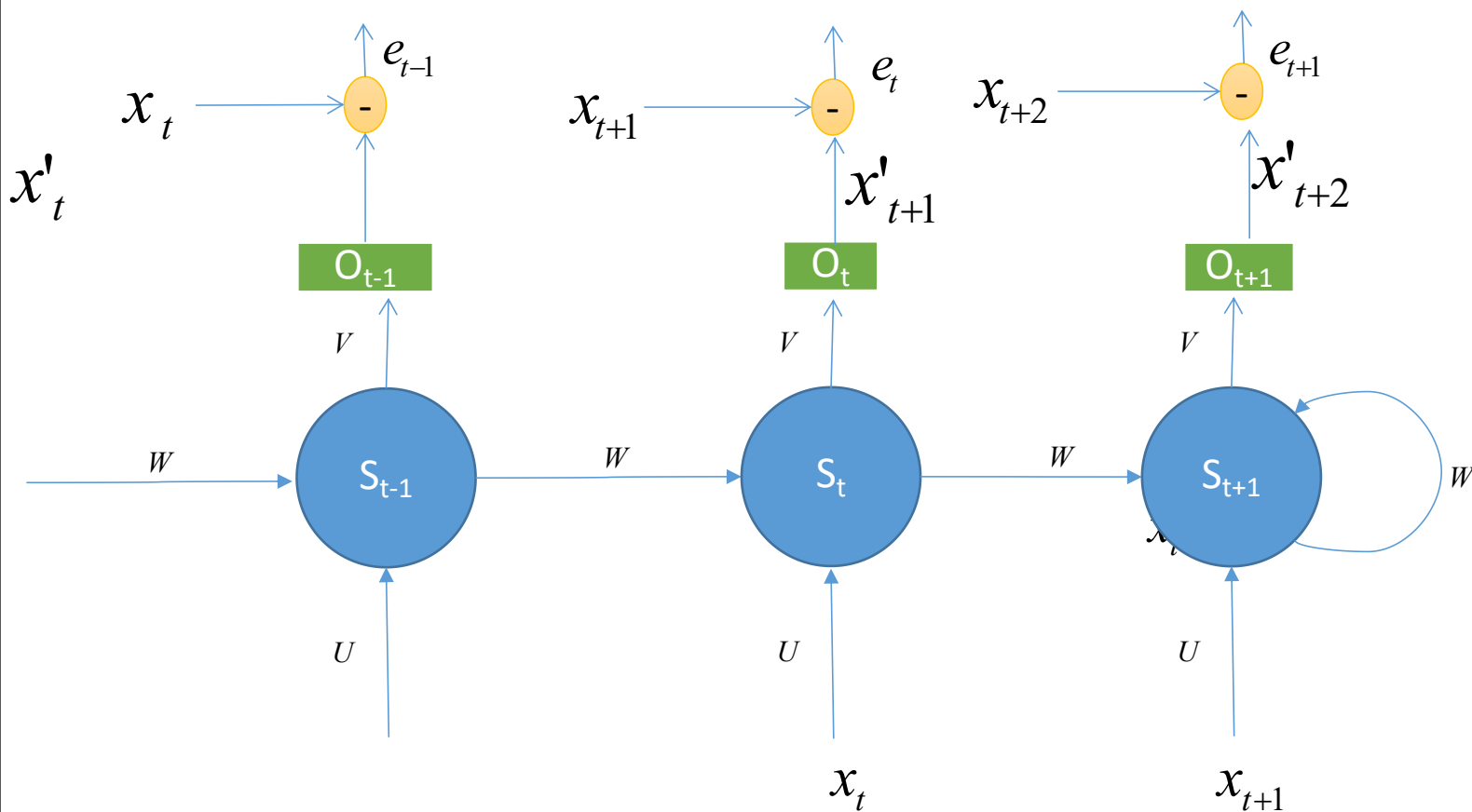
# Training Process



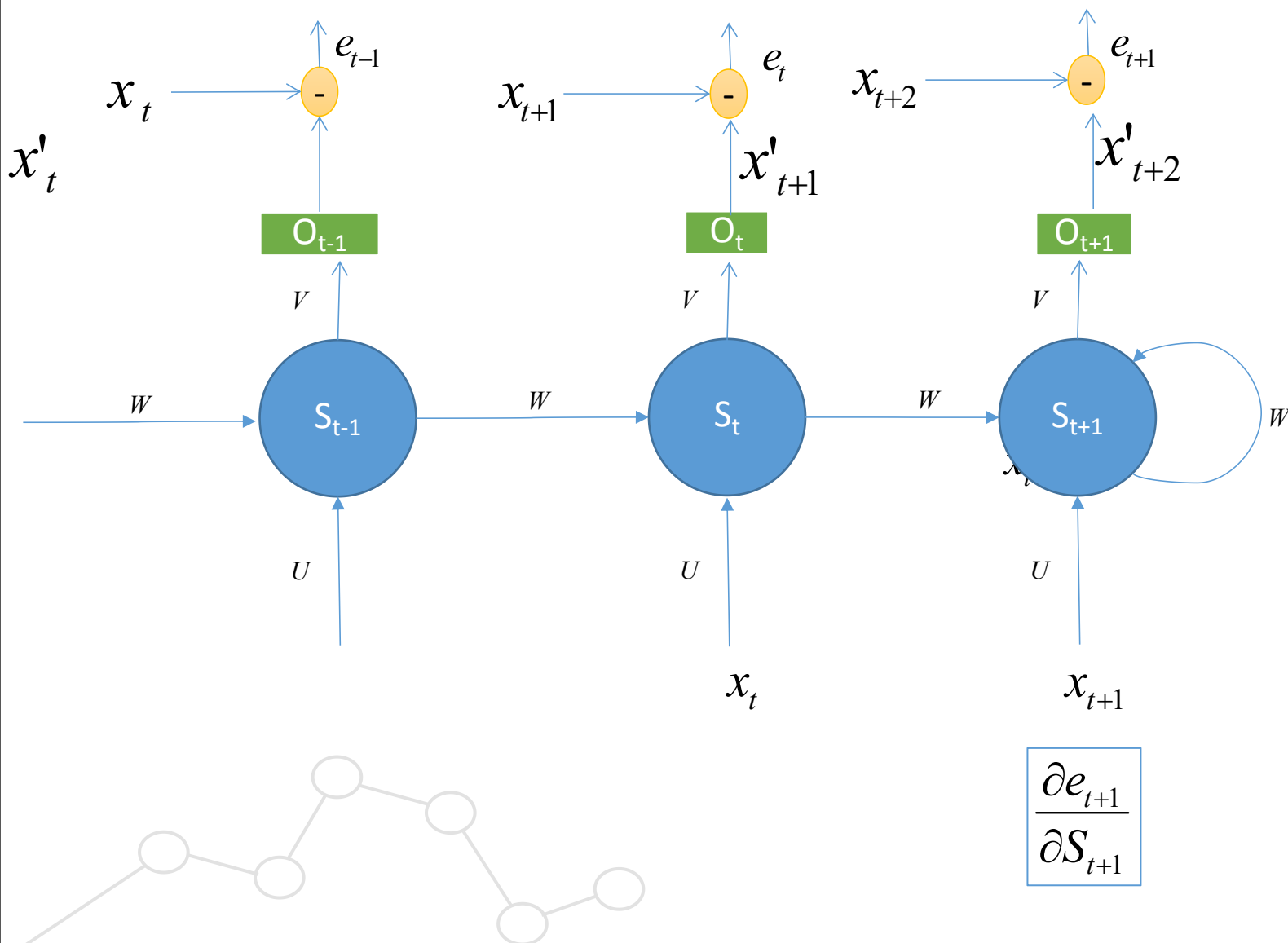
# Training Process



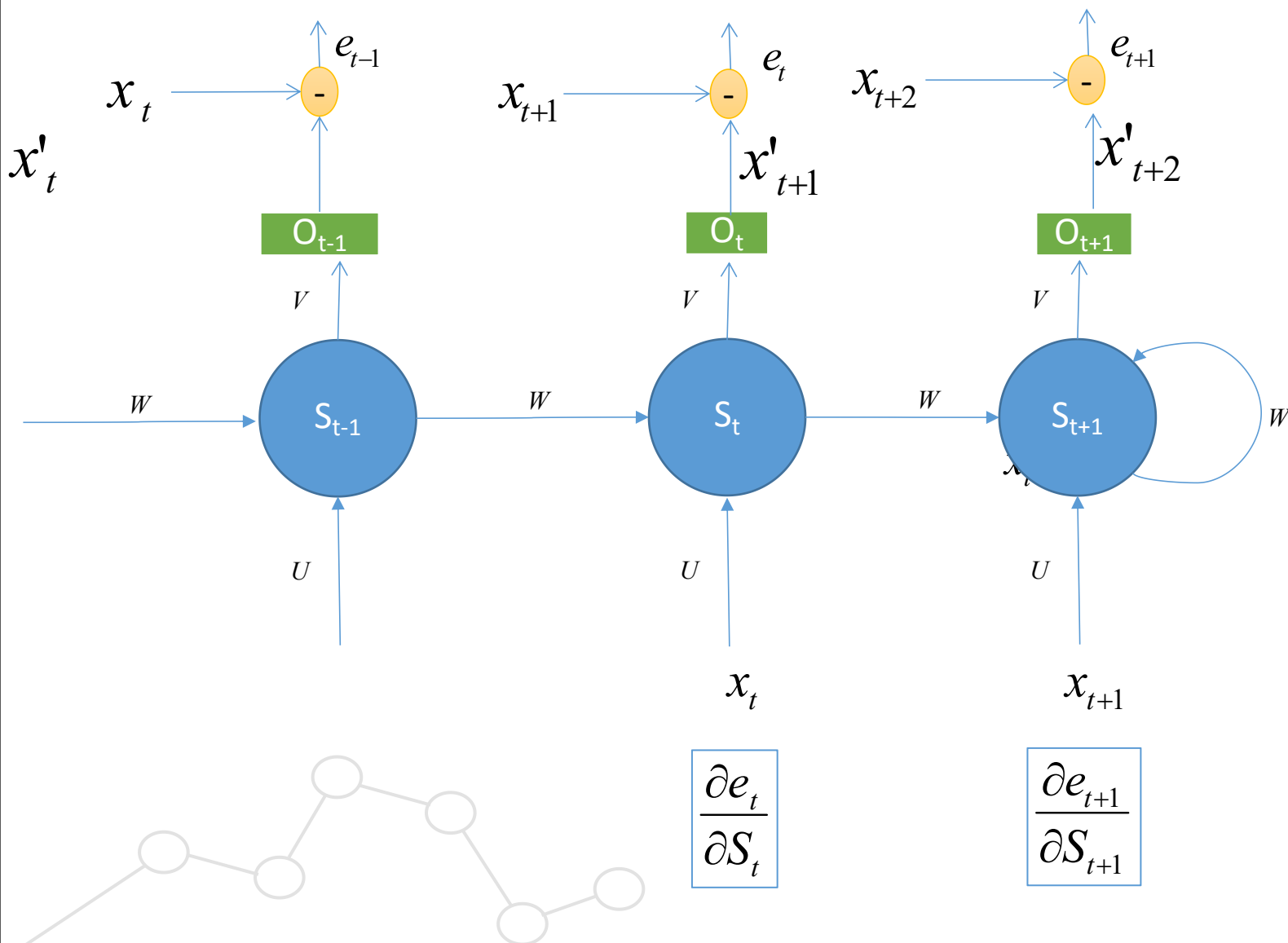
# Training Process



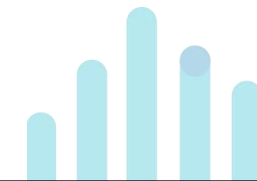
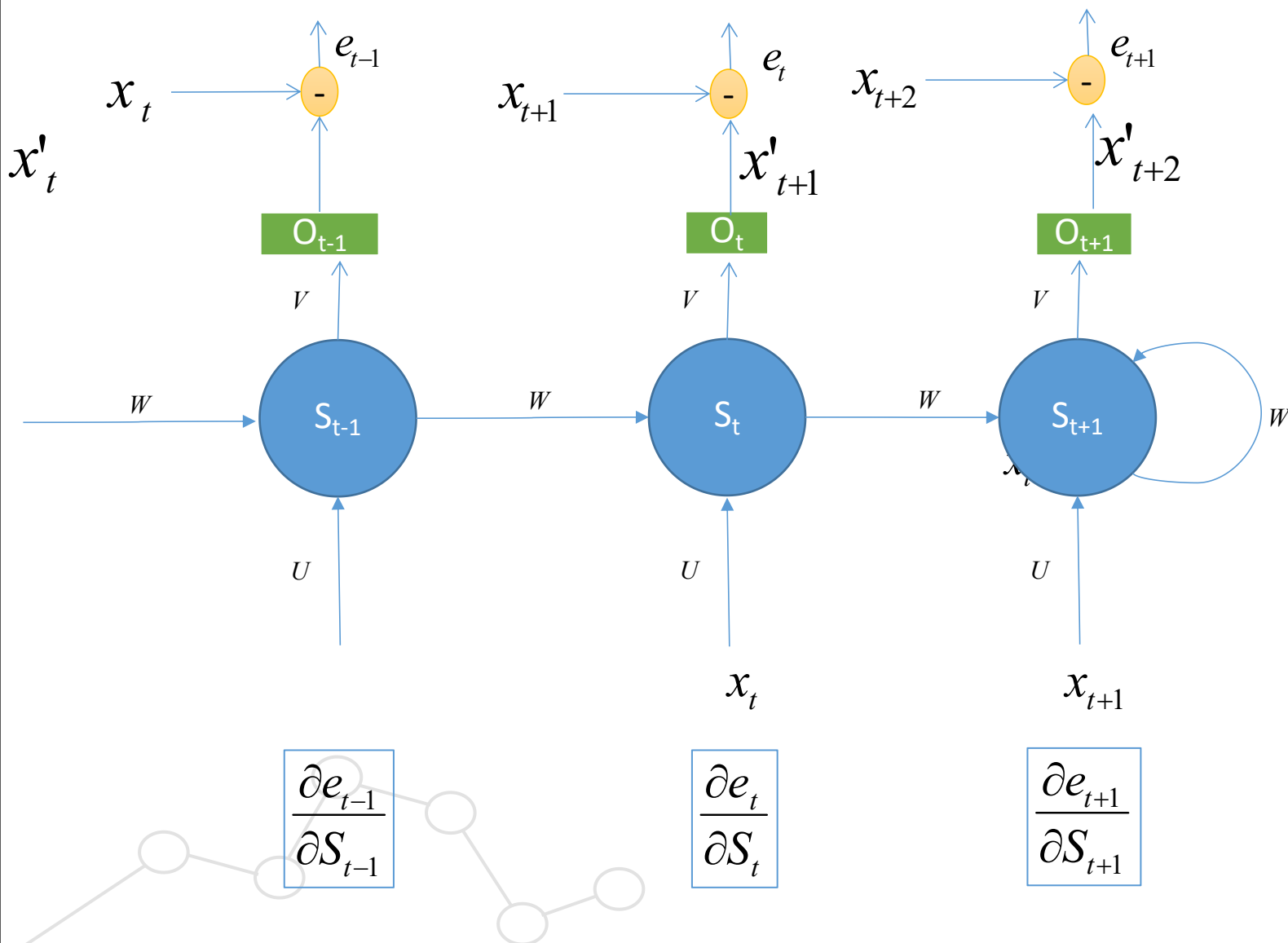
# Training Process



# Training Process

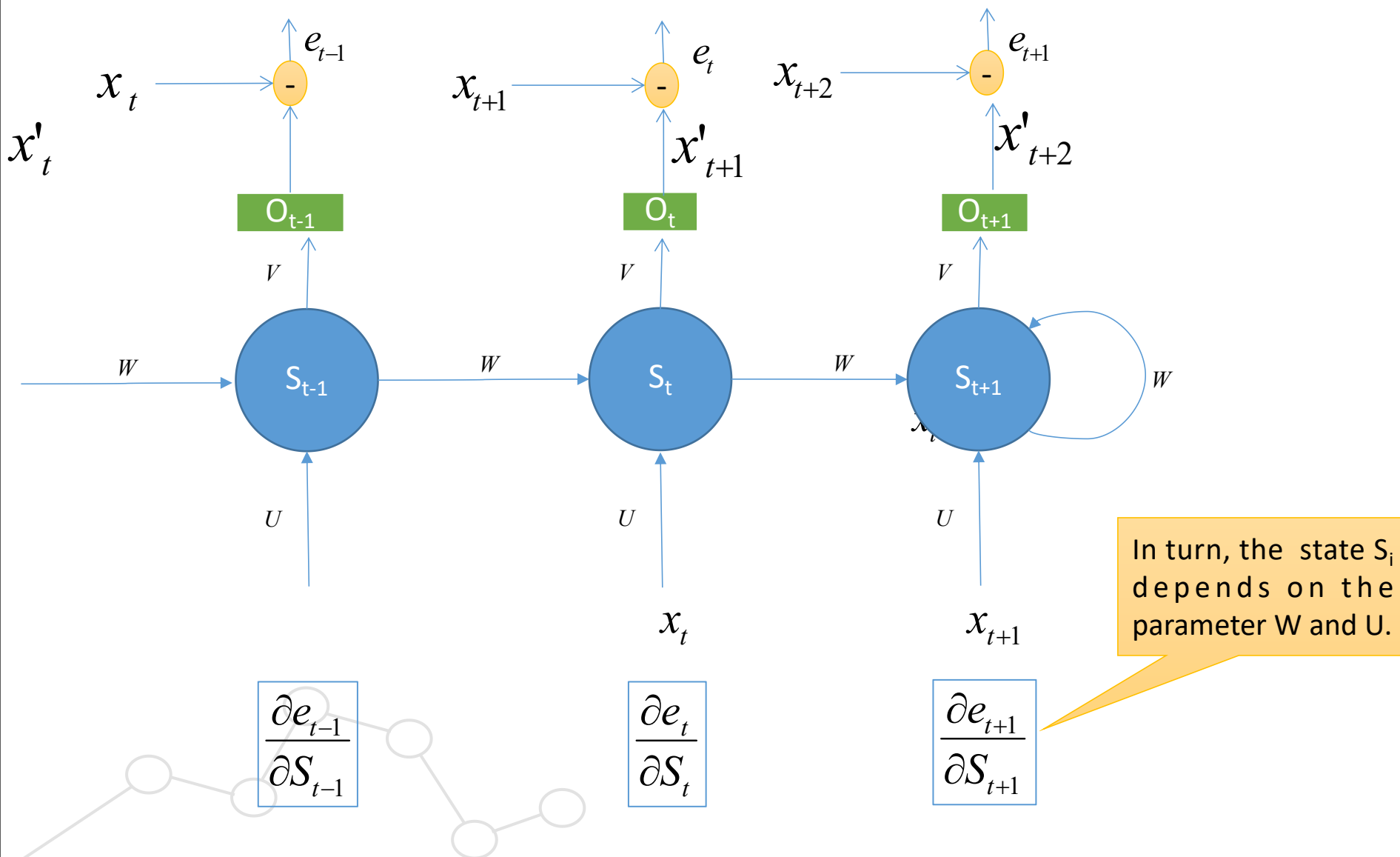


# Training Process

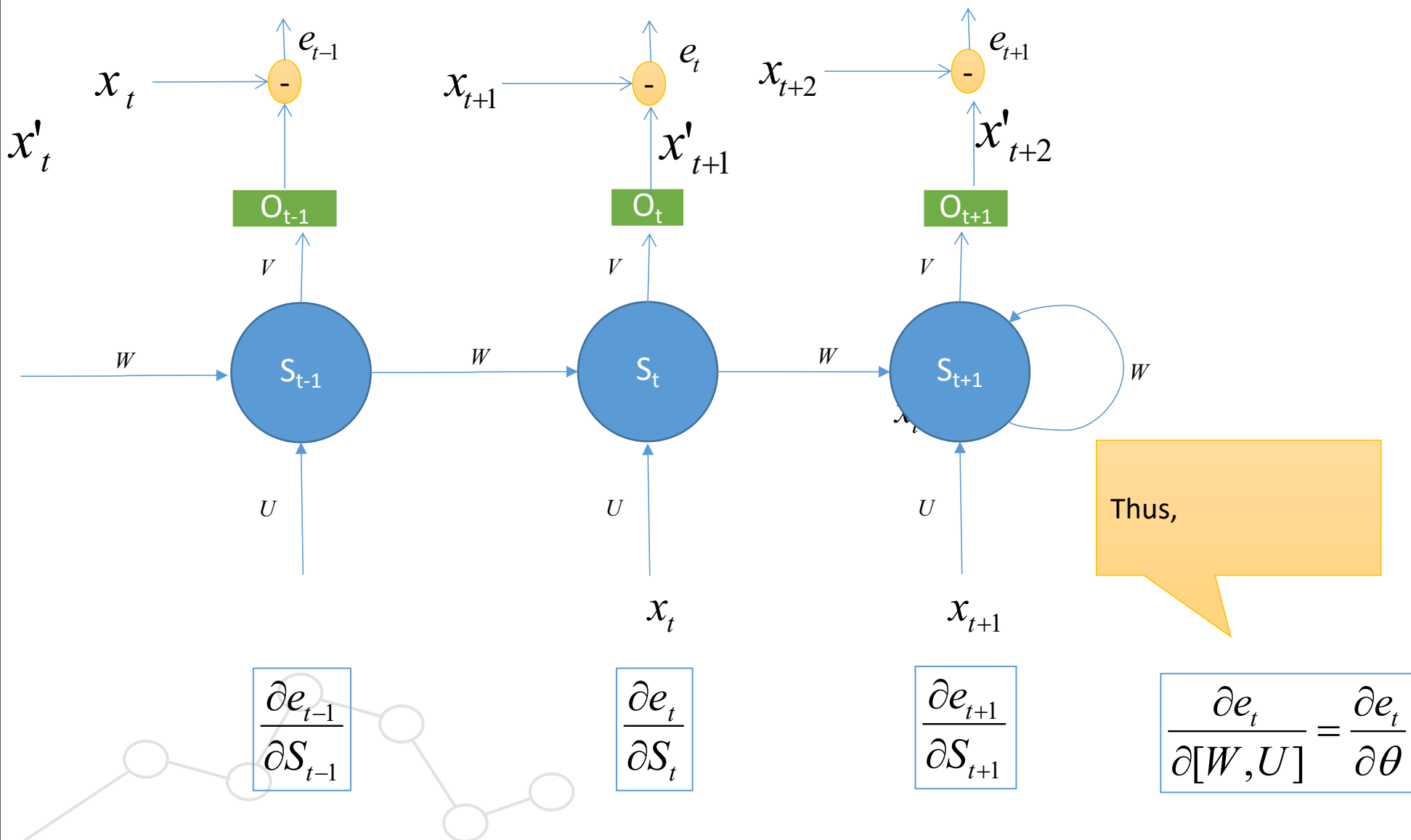




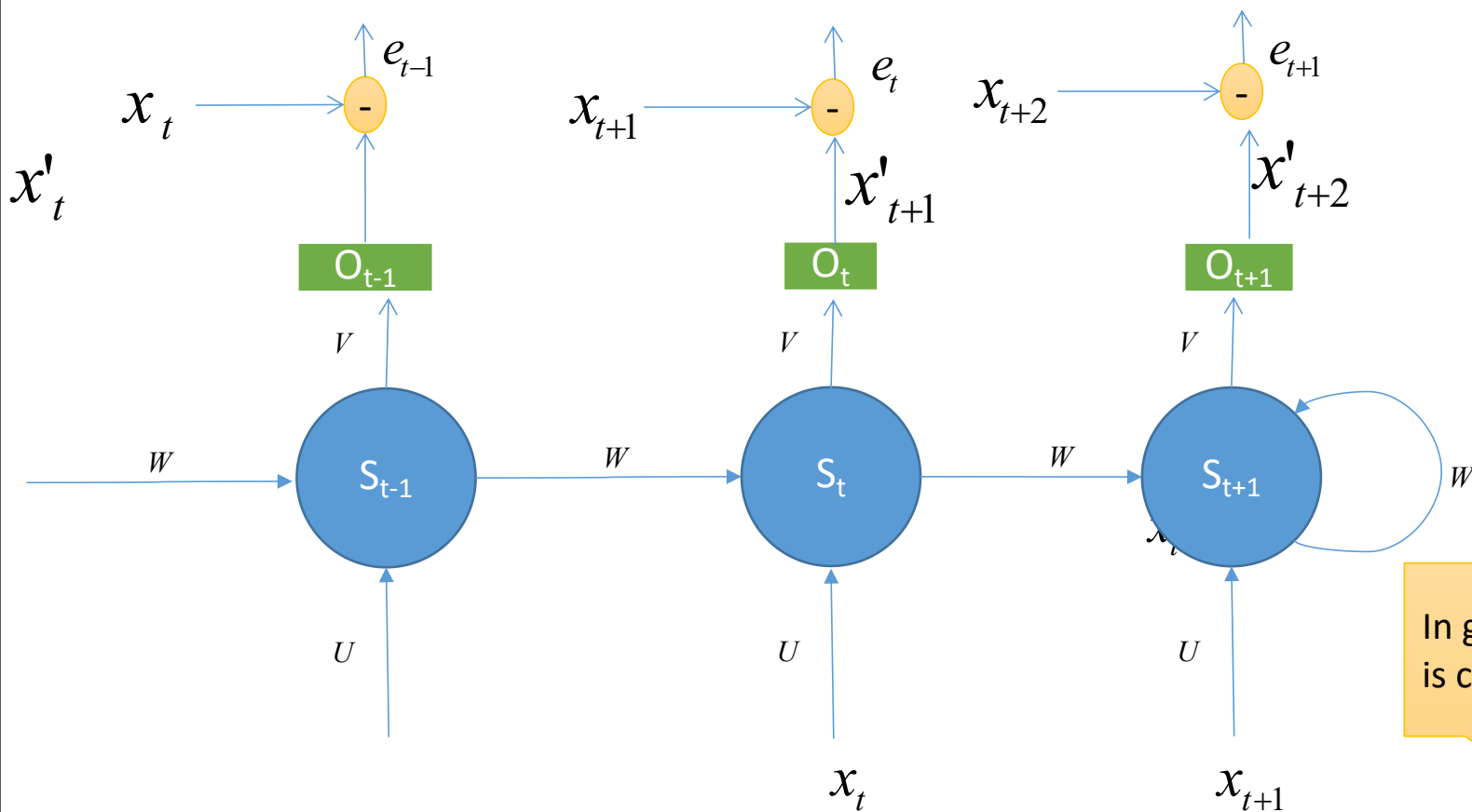
# Training Process



# Training Process



# Training Process



In general, the error is calculated as:

$$\frac{\partial e_{t-1}}{\partial S_{t-1}}$$

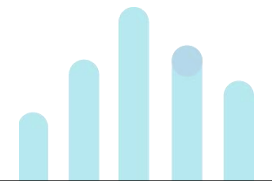
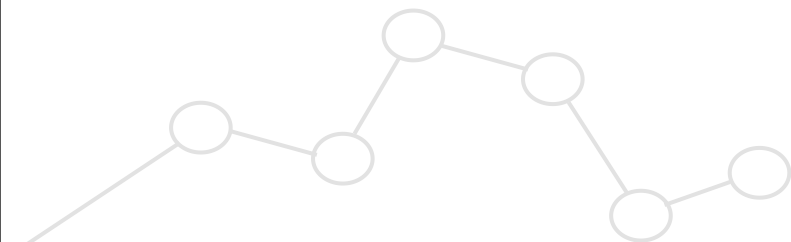
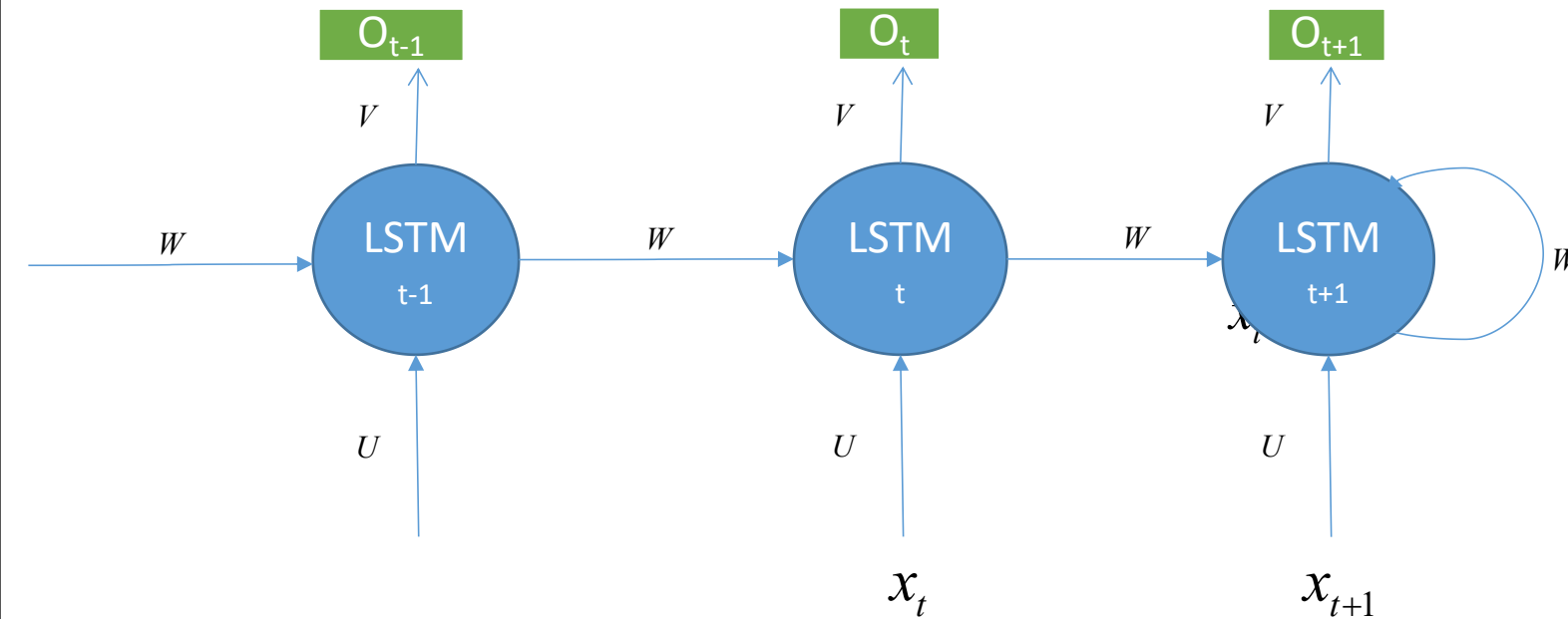
$$\frac{\partial e_t}{\partial S_t}$$

$$\frac{\partial e_{t+1}}{\partial S_{t+1}}$$

$$\frac{\partial E}{\partial \theta} = \sum_{1 \leq t \leq T} \frac{\partial e_t}{\partial \theta}$$

**Backpropagation through time BPTT**

# Long-Short Term Memory

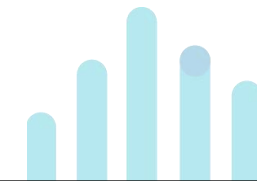
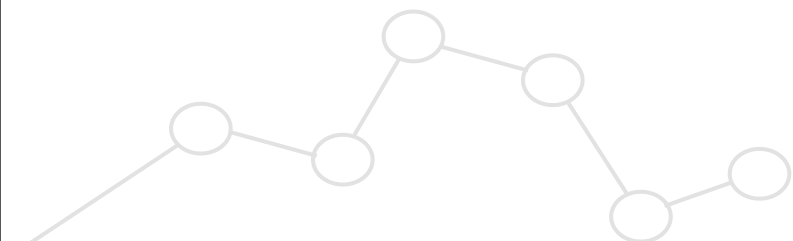


# Homework

- Review and report the main elements of the problem known as **vanishing gradient**.
- Review and report the most important elements related to LSTM (Long-Short Term Memory)
- References:

<http://proceedings.mlr.press/v28/pascanu13.pdf>

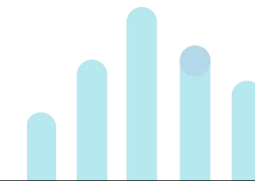
- **Deadline:** 4 de agosto de 2020



# Recurrent Neural Networks-Example

Un viejo hidalgo de los de lanza en astillero ...

Un	viejo	hidalgo	de	los	de	lanza	en	astillero
----	-------	---------	----	-----	----	-------	----	-----------



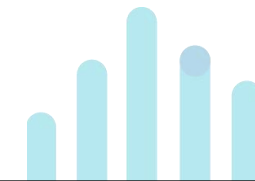
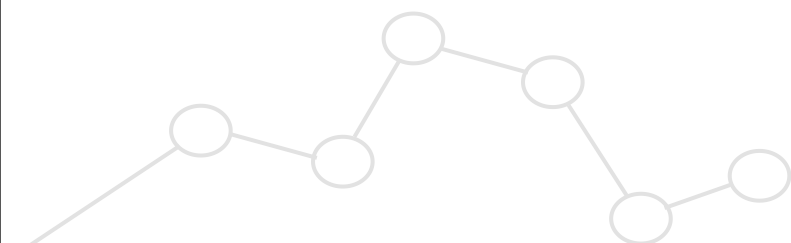
# Recurrent Neural Networks-Example

Un viejo hidalgo de los de lanza en astillero ...

Un	viejo	hidalgo	de	los	de	lanza	en	astillero	...
----	-------	---------	----	-----	----	-------	----	-----------	-----

5	451	321	8	321	8	212	10	85	...
---	-----	-----	---	-----	---	-----	----	----	-----

Indices



# Recurrent Neural Networks-Example

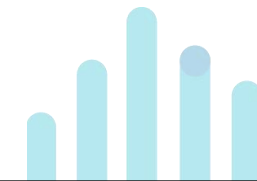
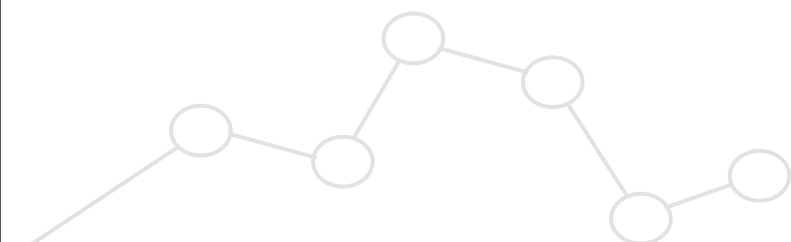
Un viejo hidalgo de los de lanza en astillero ...

Un	viejo	hidalgo	de	los	de	lanza	en	astillero	...
----	-------	---------	----	-----	----	-------	----	-----------	-----

5	451	321	8	321	8	212	10	85	...
---	-----	-----	---	-----	---	-----	----	----	-----

Indices

Size of n-gram=3





# Recurrent Neural Networks-Example

Un viejo hidalgo de los de lanza en astillero ...

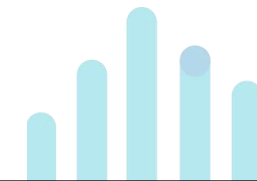
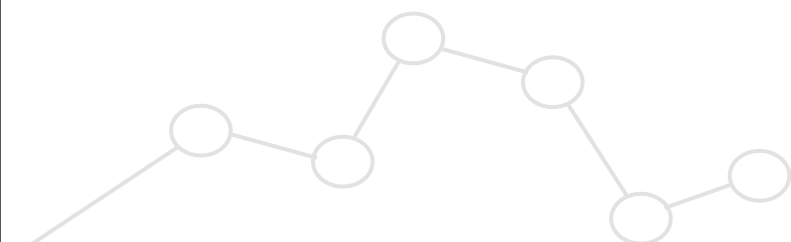
Un	viejo	hidalgo	de	los	de	lanza	en	astillero	...
----	-------	---------	----	-----	----	-------	----	-----------	-----

5	451	321	8	321	8	212	10	85	...
---	-----	-----	---	-----	---	-----	----	----	-----

Indices

Size of n-gram=3

5	451	321	8
---	-----	-----	---



# Recurrent Neural Networks-Example

Un viejo hidalgo de los de lanza en astillero ...

Un	viejo	hidalgo	de	los	de	lanza	en	astillero	...
----	-------	---------	----	-----	----	-------	----	-----------	-----

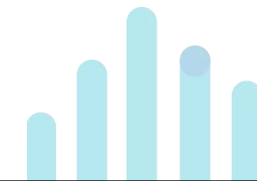
5	451	321	8	321	8	212	10	85	...
---	-----	-----	---	-----	---	-----	----	----	-----

Indices

Size of n-gram=3

5	451	321	8
---	-----	-----	---

451	321	8	321
-----	-----	---	-----



# Recurrent Neural Networks-Example

Un viejo hidalgo de los de lanza en astillero ...

Un	viejo	hidalgo	de	los	de	lanza	en	astillero	...
----	-------	---------	----	-----	----	-------	----	-----------	-----

5	451	321	8	321	8	212	10	85	...
---	-----	-----	---	-----	---	-----	----	----	-----

Indices

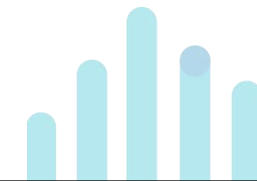
Size of n-gram=3

5	451	321	8
---	-----	-----	---

451	321	8	321
-----	-----	---	-----

321	8	321	8
-----	---	-----	---

Labeled dataset



# Recurrent Neural Networks-Example

Un viejo hidalgo de los de lanza en astillero ...

Un	viejo	hidalgo	de	los	de	lanza	en	astillero	...
----	-------	---------	----	-----	----	-------	----	-----------	-----

5	451	321	8	321	8	212	10	85	...
---	-----	-----	---	-----	---	-----	----	----	-----

Indices

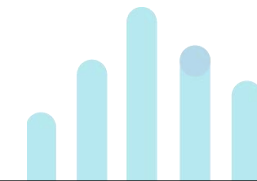
Size of n-gram=3

5	451	321	8
---	-----	-----	---

451	321	8	321
-----	-----	---	-----

321	8	321	8
-----	---	-----	---

N= number of sequences



# Recurrent Neural Networks-Example

Un viejo hidalgo de los de lanza en astillero ...

Un	viejo	hidalgo	de	los	de	lanza	en	astillero	...
----	-------	---------	----	-----	----	-------	----	-----------	-----

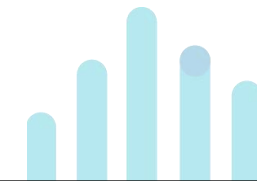
5	451	321	8	321	8	212	10	85	...
---	-----	-----	---	-----	---	-----	----	----	-----

Indices

Size of n-gram=3



N= number of sequences  
Size of window=10



# Recurrent Neural Networks-Example

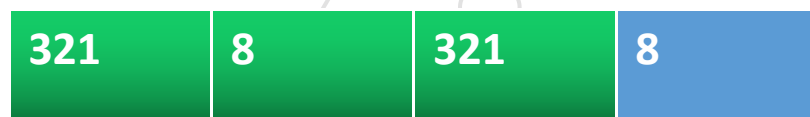
Un viejo hidalgo de los de lanza en astillero ...

Un	viejo	hidalgo	de	los	de	lanza	en	astillero	...
----	-------	---------	----	-----	----	-------	----	-----------	-----

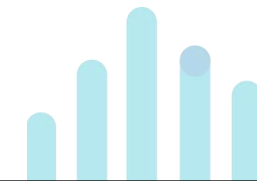
5	451	321	8	321	8	212	10	85	...
---	-----	-----	---	-----	---	-----	----	----	-----

Indices

Size of n-gram=3



N= number of sequences  
Size of window=10



# Recurrent Neural Networks-Example

Un viejo hidalgo de los de lanza en astillero ...

Un	viejo	hidalgo	de	los	de	lanza	en	astillero	...
----	-------	---------	----	-----	----	-------	----	-----------	-----

5	451	321	8	321	8	212	10	85	...
---	-----	-----	---	-----	---	-----	----	----	-----

Indices

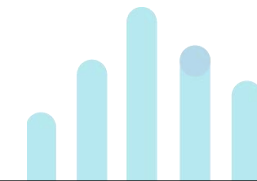
Size of n-gram=3



451	321	8	321
-----	-----	---	-----

321	8	321	8
-----	---	-----	---

N= number of sequences  
Size of window=10



# Recurrent Neural Networks-Example

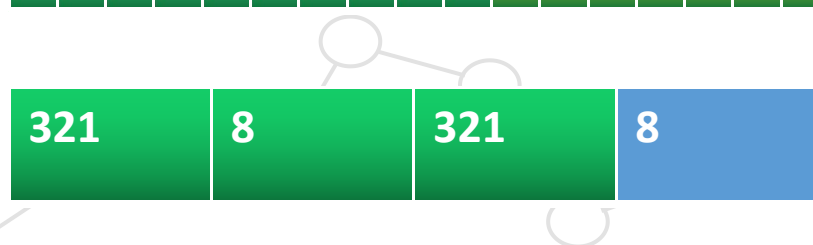
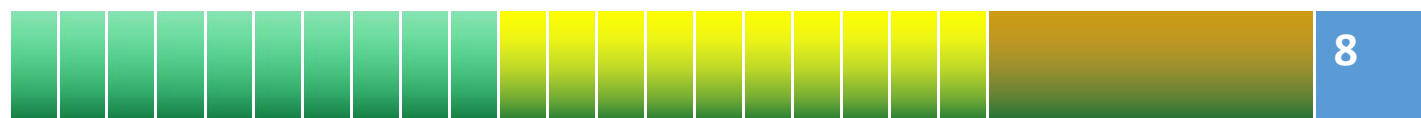
Un viejo hidalgo de los de lanza en astillero ...

Un	viejo	hidalgo	de	los	de	lanza	en	astillero	...
----	-------	---------	----	-----	----	-------	----	-----------	-----

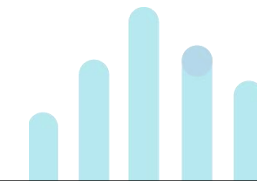
5	451	321	8	321	8	212	10	85	...
---	-----	-----	---	-----	---	-----	----	----	-----

Indices

Size of n-gram=3



N= number of sequences  
Size of window=10





# Recurrent Neural Networks-Example

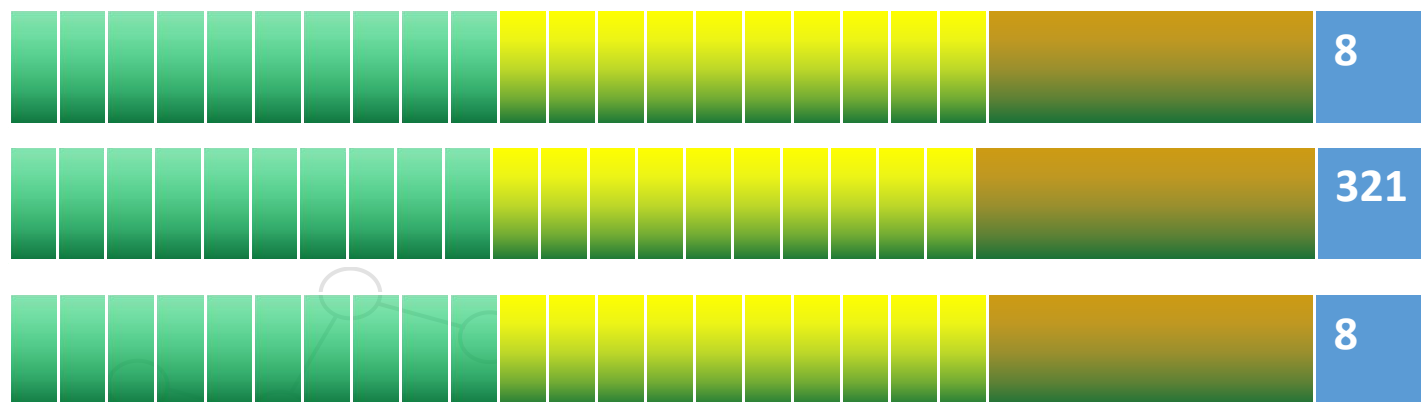
Un viejo hidalgo de los de lanza en astillero ...

Un	viejo	hidalgo	de	los	de	lanza	en	astillero	...
----	-------	---------	----	-----	----	-------	----	-----------	-----

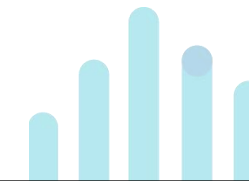
5	451	321	8	321	8	212	10	85	...
---	-----	-----	---	-----	---	-----	----	----	-----

Indices

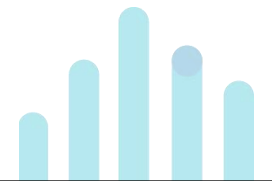
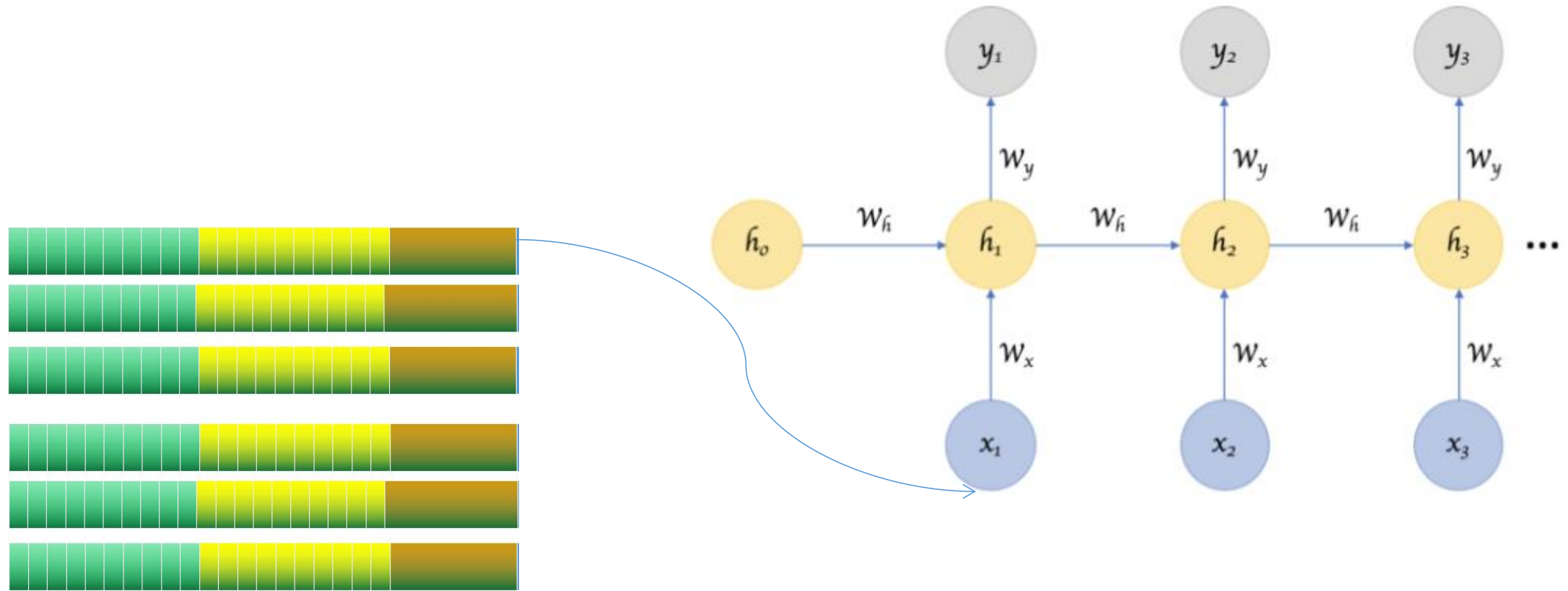
Size of n-gram=3



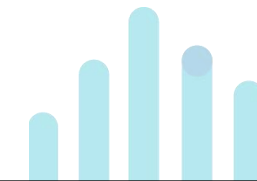
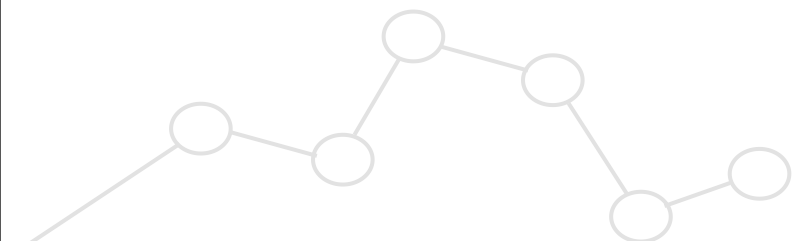
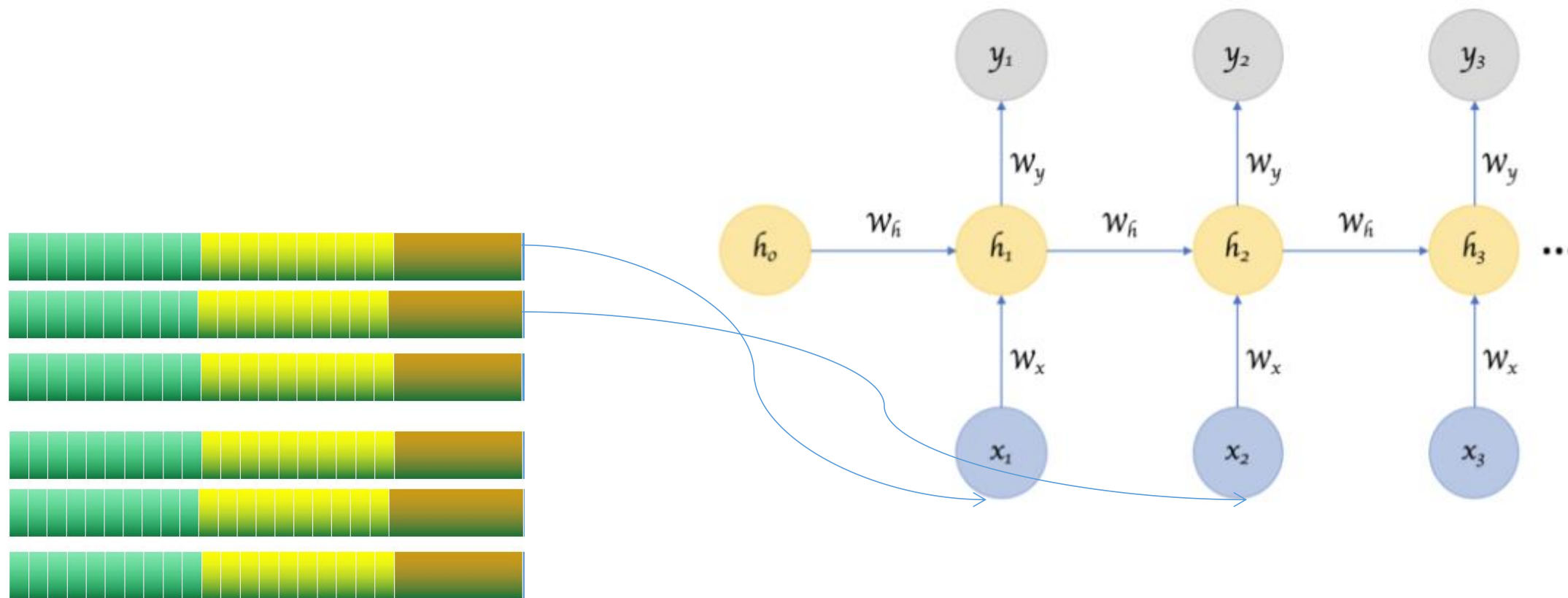
N= number of sequences  
Size of window=10



# Recurrent Neural Networks-Example



# Recurrent Neural Networks-Example



# Recurrent Neural Networks-Example

