DATA SCIENCE

DATE : 14.10.2024

DT/NT : DT

LESSON : DEEP LEARNING

SUBJECT: RECURRENT NEURAL NETWORK – LSTM - GRU

BATCH : 250

TECHPRO EDUCATION

techproeducation.com

+1 (585) 304 29 59

DEEP LEARNING II

Recurrent Neural Network(RNN)

GENERAL CONTENTS

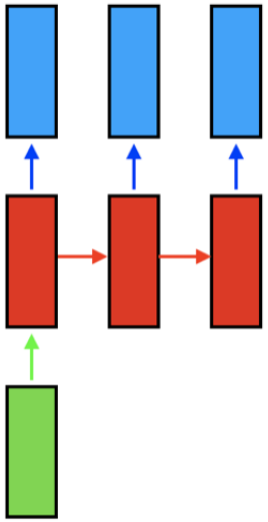
- Recurrent Neural Network (RNN)
- Long Short-Term Memory (LSTM)
- Gated Recurrent Units (GRU)

Machine Learning

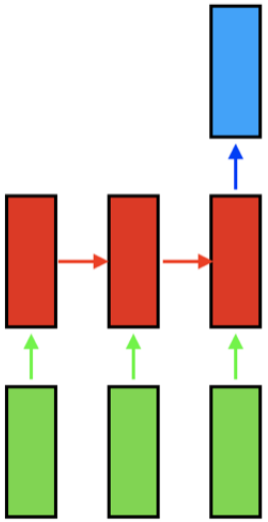
Different Types of Sequence Modeling Tasks

one to one

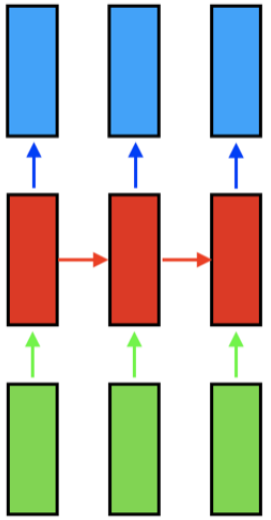
one to many



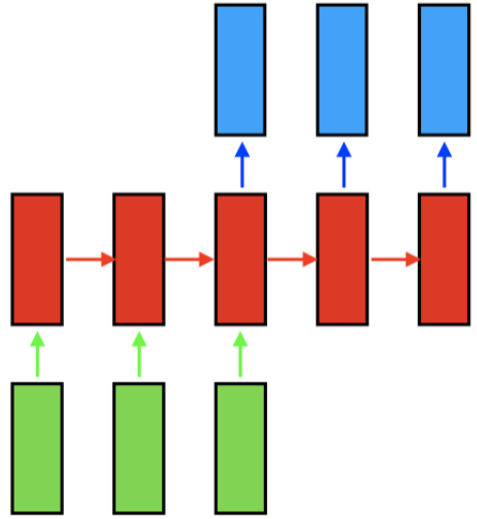
many to one



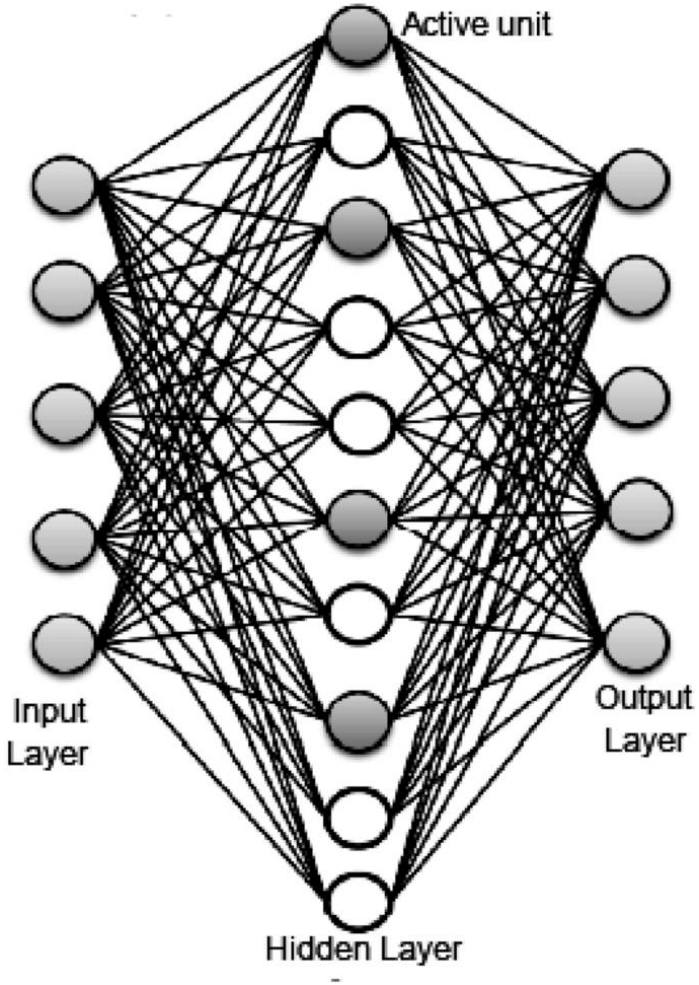
many to many

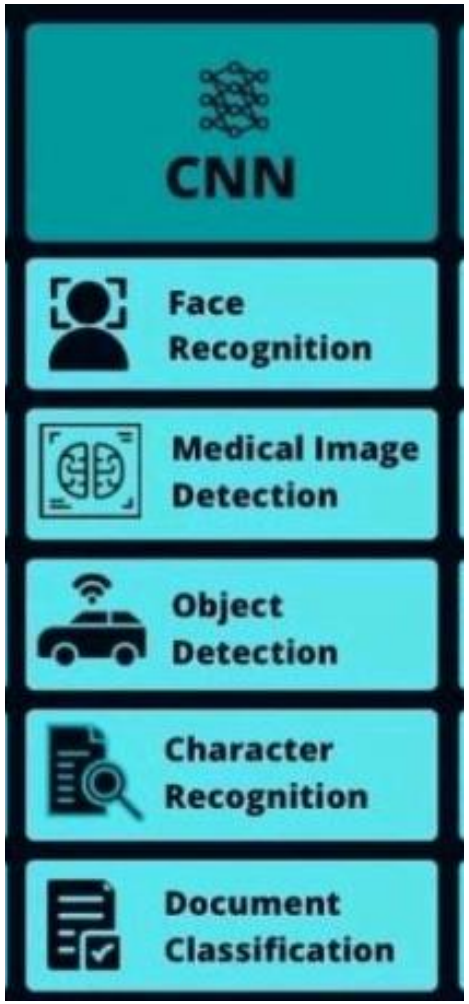


many to many



RNN

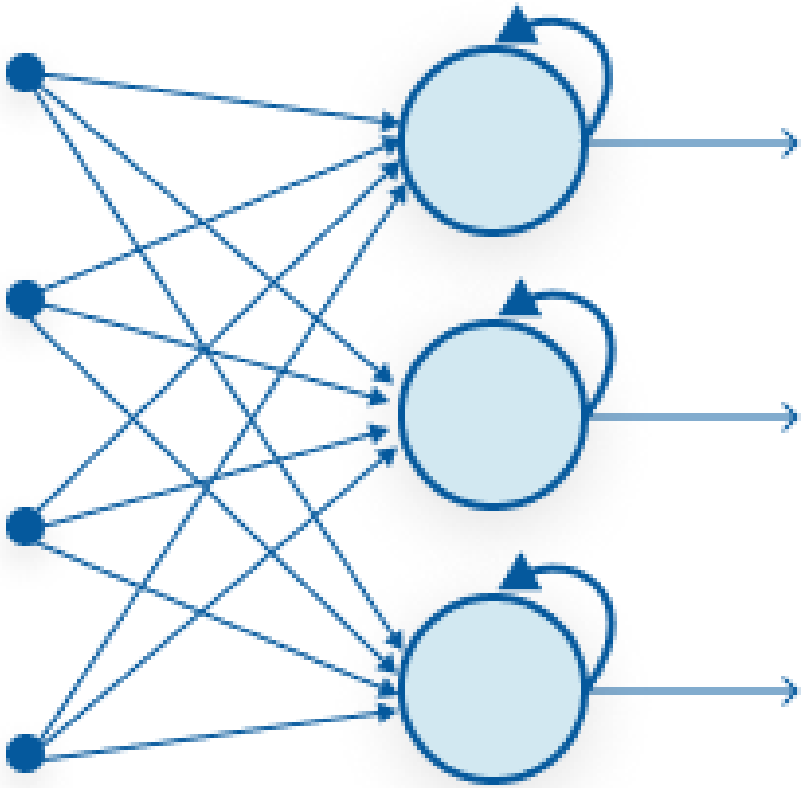

Active unit
Input Layer
Output Layer
Hidden Layer


CNN
Face Recognition
Medical Image Detection
Object Detection
Character Recognition
Document Classification

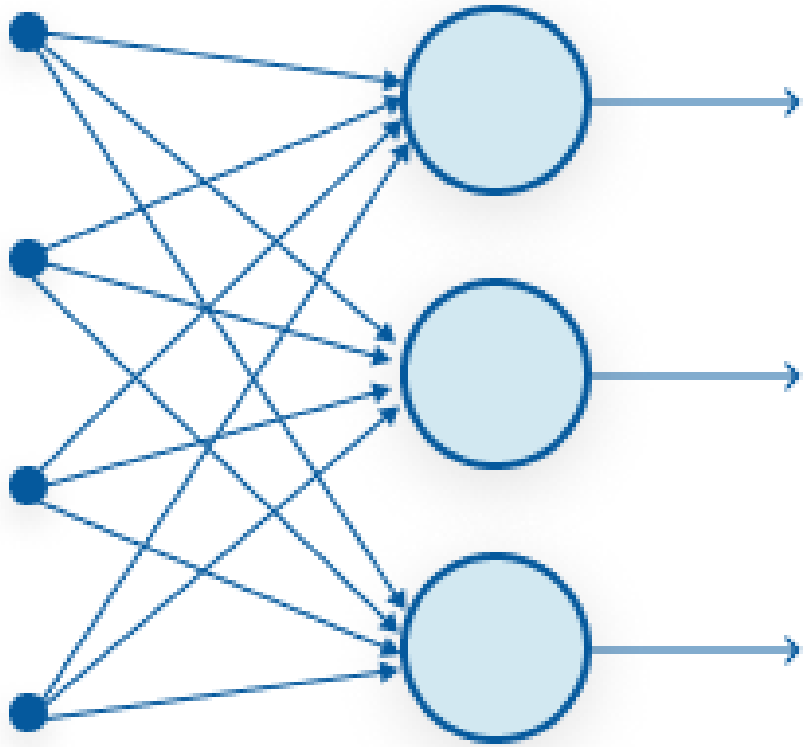

RNN
Stock Price Prediction
Chatbots
Voice Assistant
Weather Forecasting
Music Generation

DIFFERENCE BETWEEN RNN AND ANN

RNN



Recurrent Neural Network



Feed-Forward Neural Network

RNN

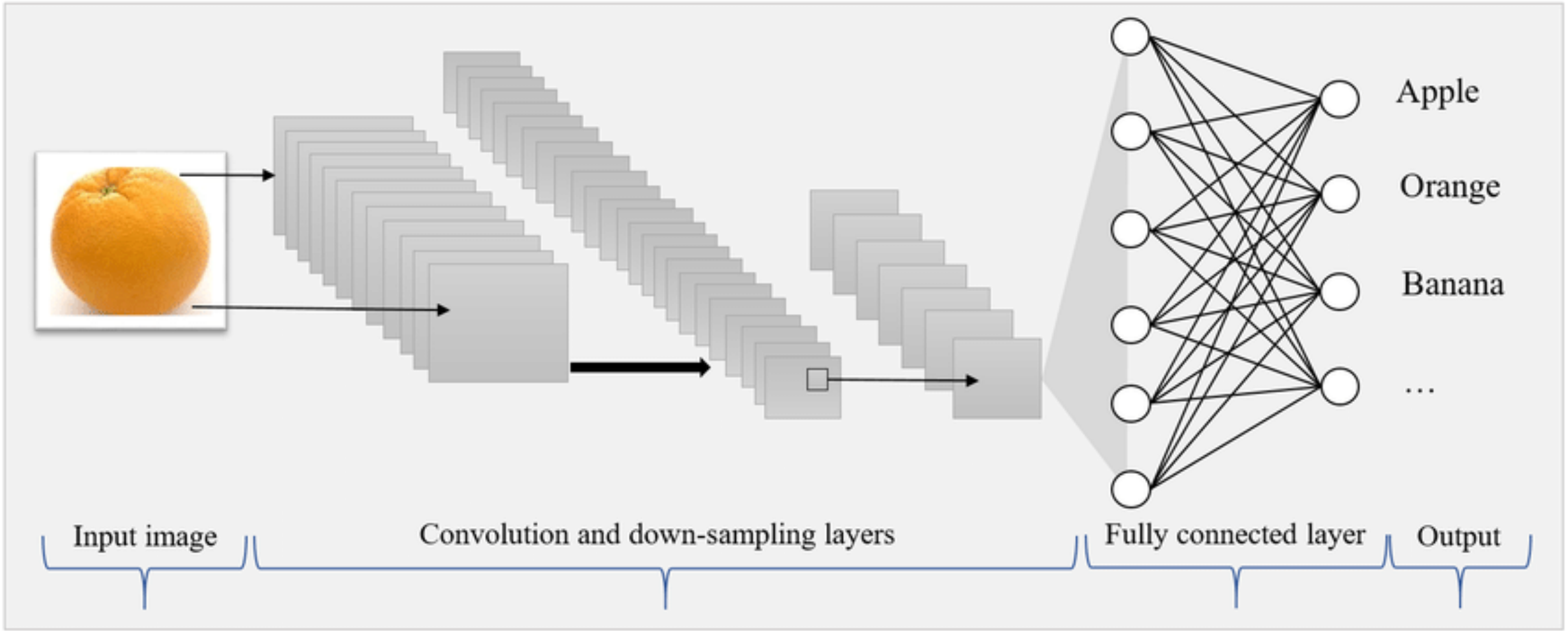
ANN

01 cannot handle sequential data

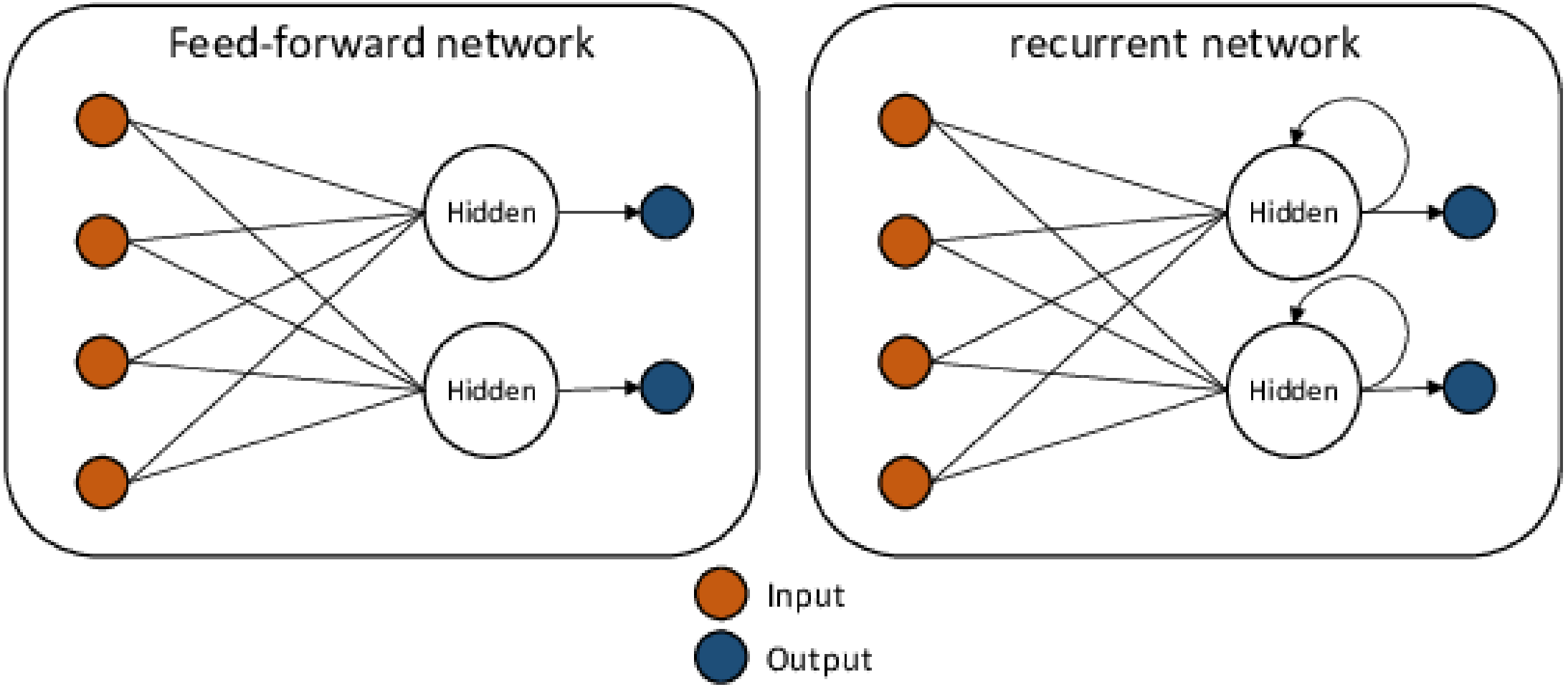
02 considers only the current input

03 cannot memorize previous inputs

RNN


Apple
Orange
Banana
...
Input image
Convolution and down-sampling layers
Fully connected layer
Output

RNN


Feed-forward network
Hidden
Hidden
recurrent network
Hidden
Hidden
Input
Output

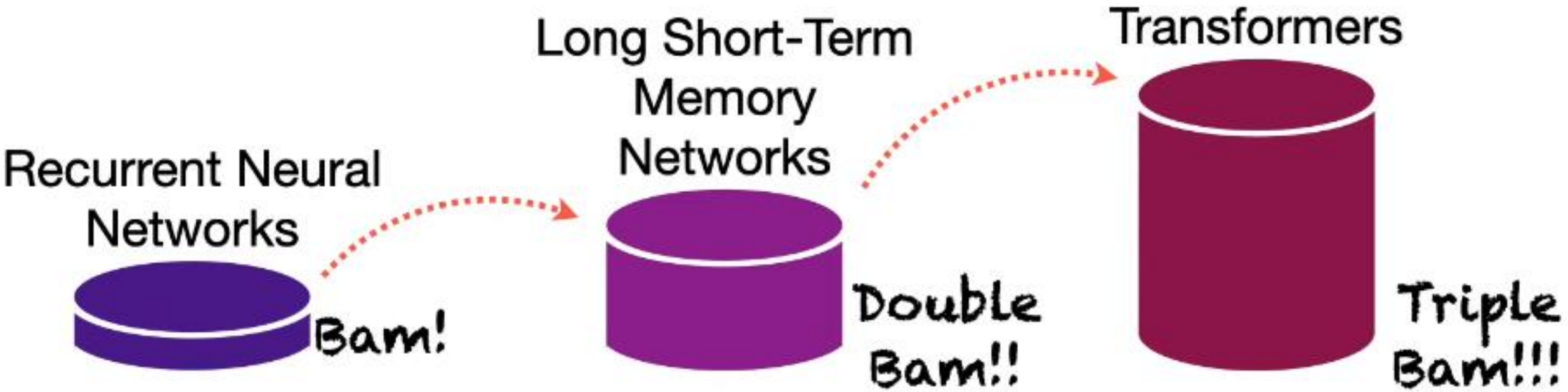
RNN

- RNNs are preferred in NLP because they are effective at modeling sequential data.
- The key difference between RNN models and ANN or CNN is that each layer in an RNN depends on the information from the previous layer. Without this information from the previous layer, it cannot function.
- In CNNs and ANNs, all data is processed independently of each other.
- Another key difference of RNN compared to CNN and ANN is that RNN has memory, which is why it is often preferred in NLP models.

Vanilla-RNN

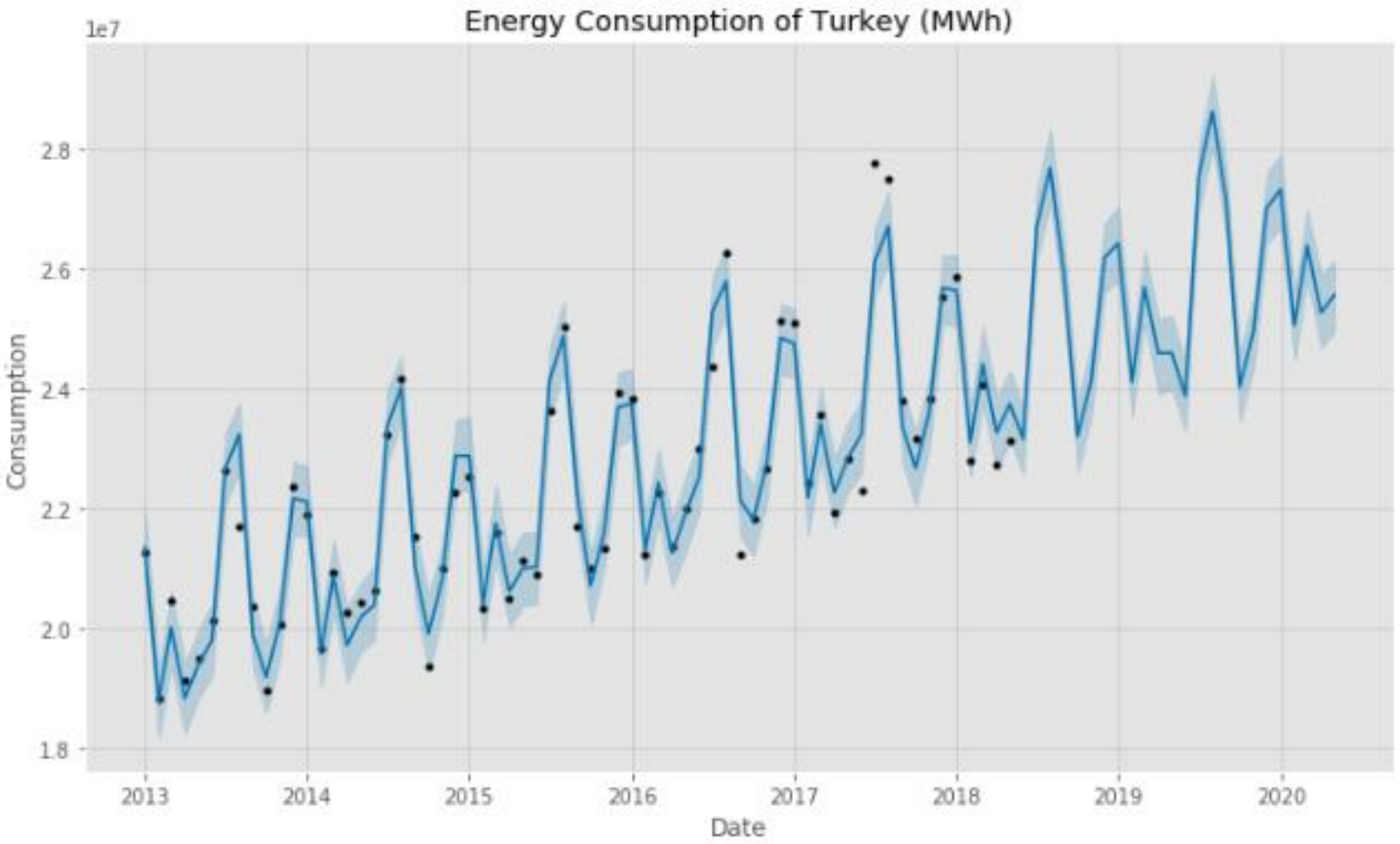
RNN (Recurrent Neural Network)

RNN



Source: StatQuest with Josh Starmer

RNN



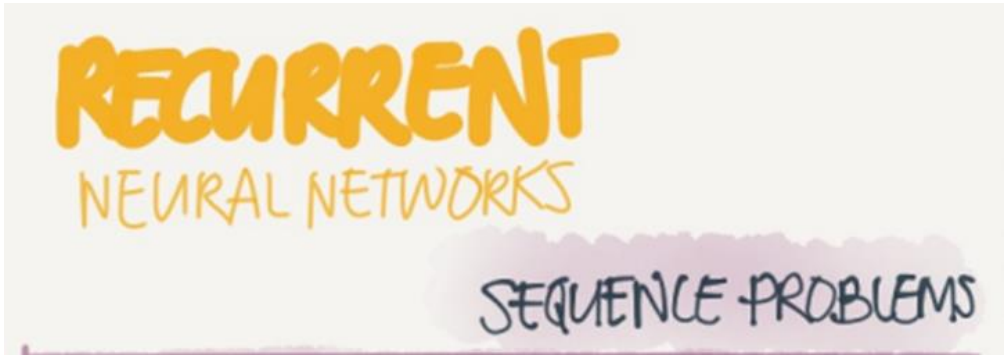
Time Series Example

	A	B
1	date	value
2	9/6/2017	531974.19
3	9/7/2017	484704.26
4	9/8/2017	693635.27
5	9/9/2017	420176.65
6	9/10/2017	257548.74
7	9/11/2017	212416.06
8	9/12/2017	410240.57
9	9/13/2017	559267.26
10	9/14/2017	556496.67
11	9/15/2017	813277.37
12	9/16/2017	600138.13
13	9/17/2017	371246.62
14	9/18/2017	319319.61
15	9/19/2017	561685.94
16	9/20/2017	650536.61
17	9/21/2017	599229.88

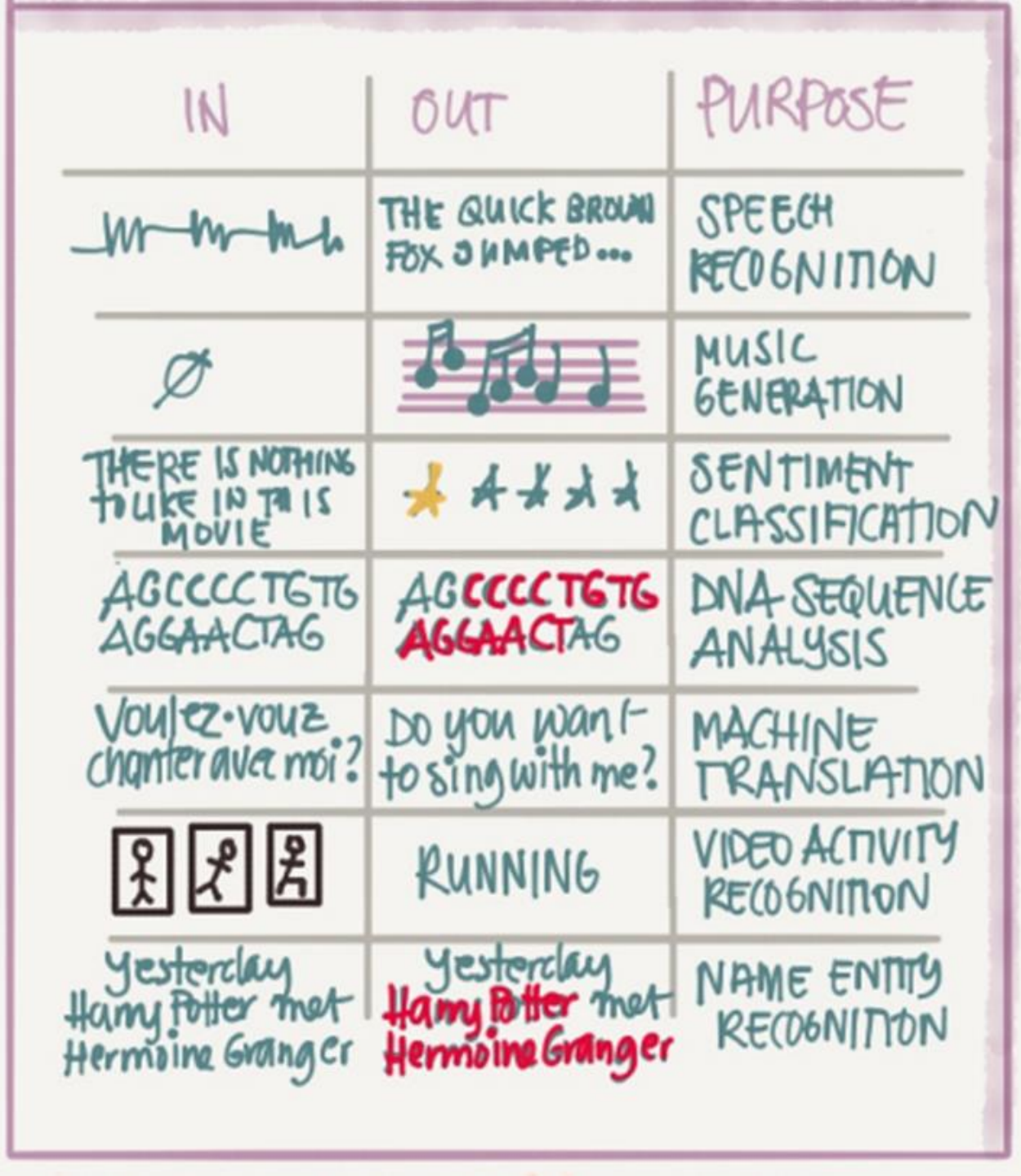
Time Series Example

THE USE CASES OF RNN IN REAL LIFE

RNN



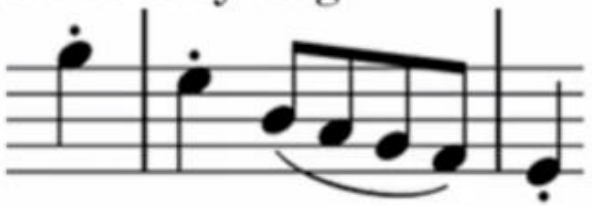
A Timeseries is a common example of this, with each point reflecting an observation at a certain point in time, such as a stock price or sensor data.



Examples of sequence data

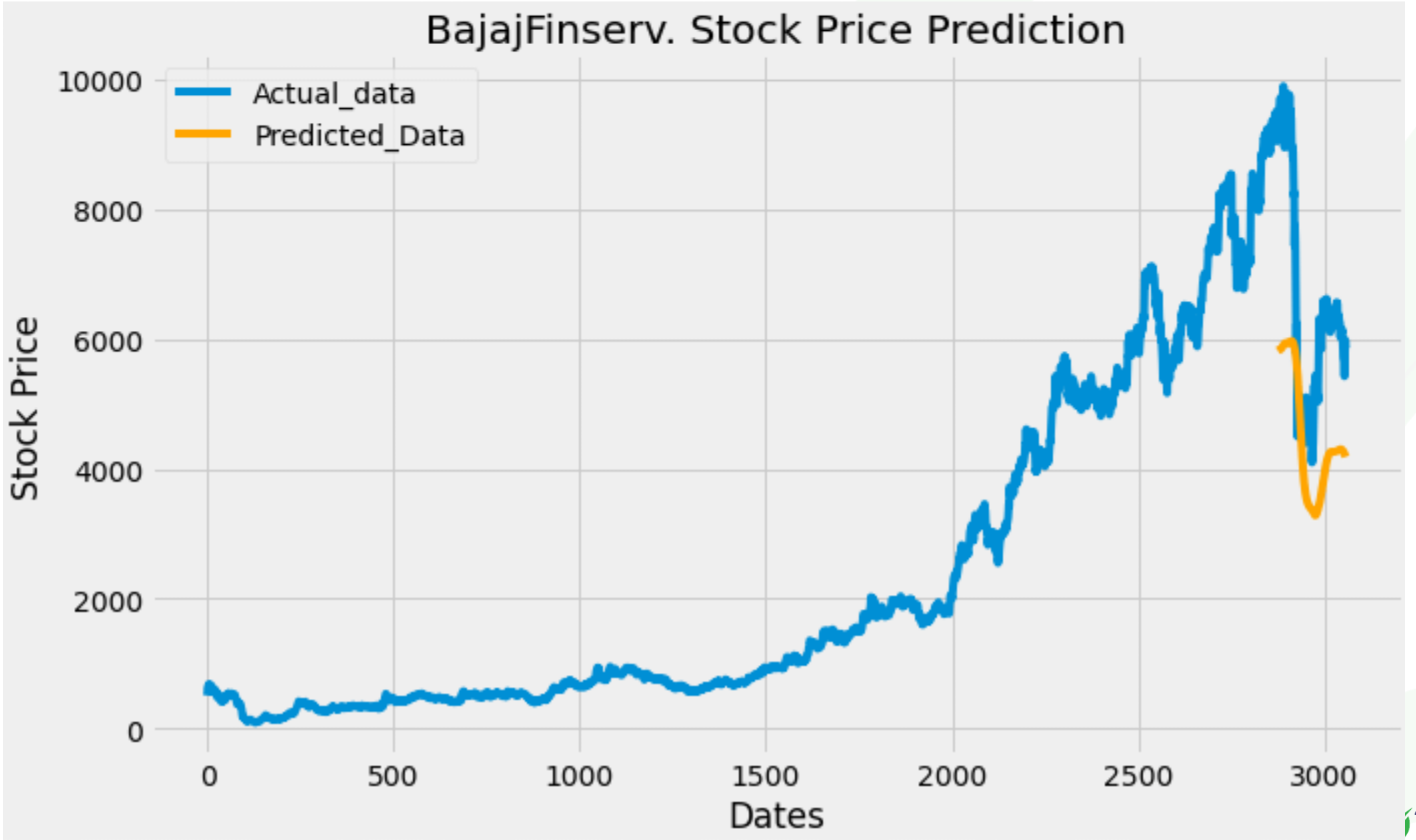
x

y

Speech recognition		→	"The quick brown fox jumped over the lazy dog."
Music generation	∅	→	
Sentiment classification	"There is nothing to like in this movie."	→	★☆☆☆☆
DNA sequence analysis	AGCCCCTGTGAGGAACTAG	→	AGCCCCTGTGAGGAACTAG
Machine translation	Voulez-vous chanter avec moi?	→	Do you want to sing with me?
Video activity recognition		→	Running
Name entity recognition	Yesterday, Harry Potter met Hermione Granger.	→	Yesterday, Harry Potter met Hermione Granger.

Andrew Ng

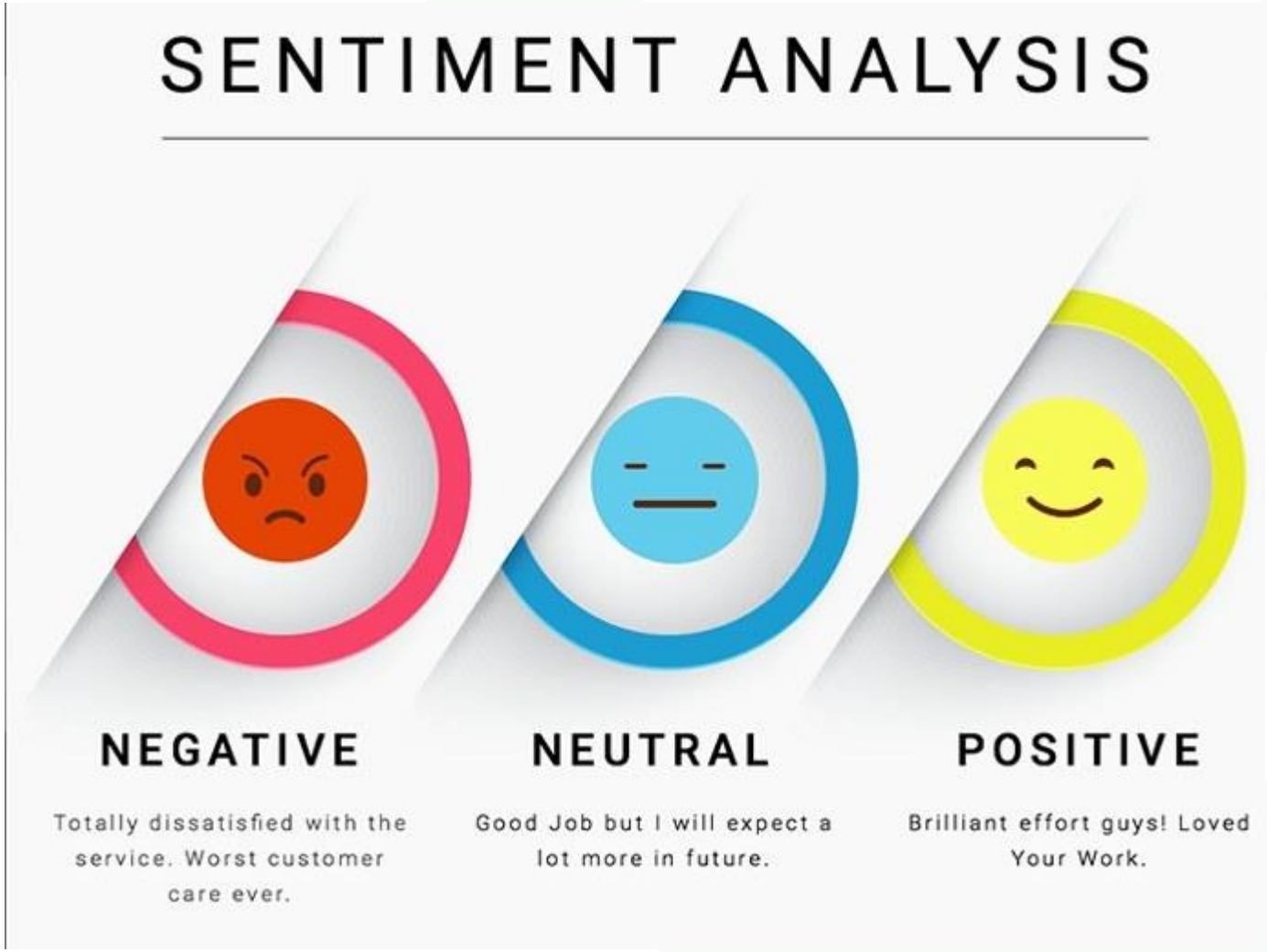
RNN


BajajFinserv. Stock Price Prediction
Actual_data
Predicted_Data
Stock Price
10000
8000
6000
4000
2000
0
0
500
1000
1500
2000
2500
3000
Dates

RNN APPLICATION

Machine Learning

Text Mining and Sentiment Analysis can be carried out using RNN for Natural Language Processing



RNN APPLICATION

Here the person is speaking in English and it is getting translated into Chinese, Italian, French, German and Spanish languages

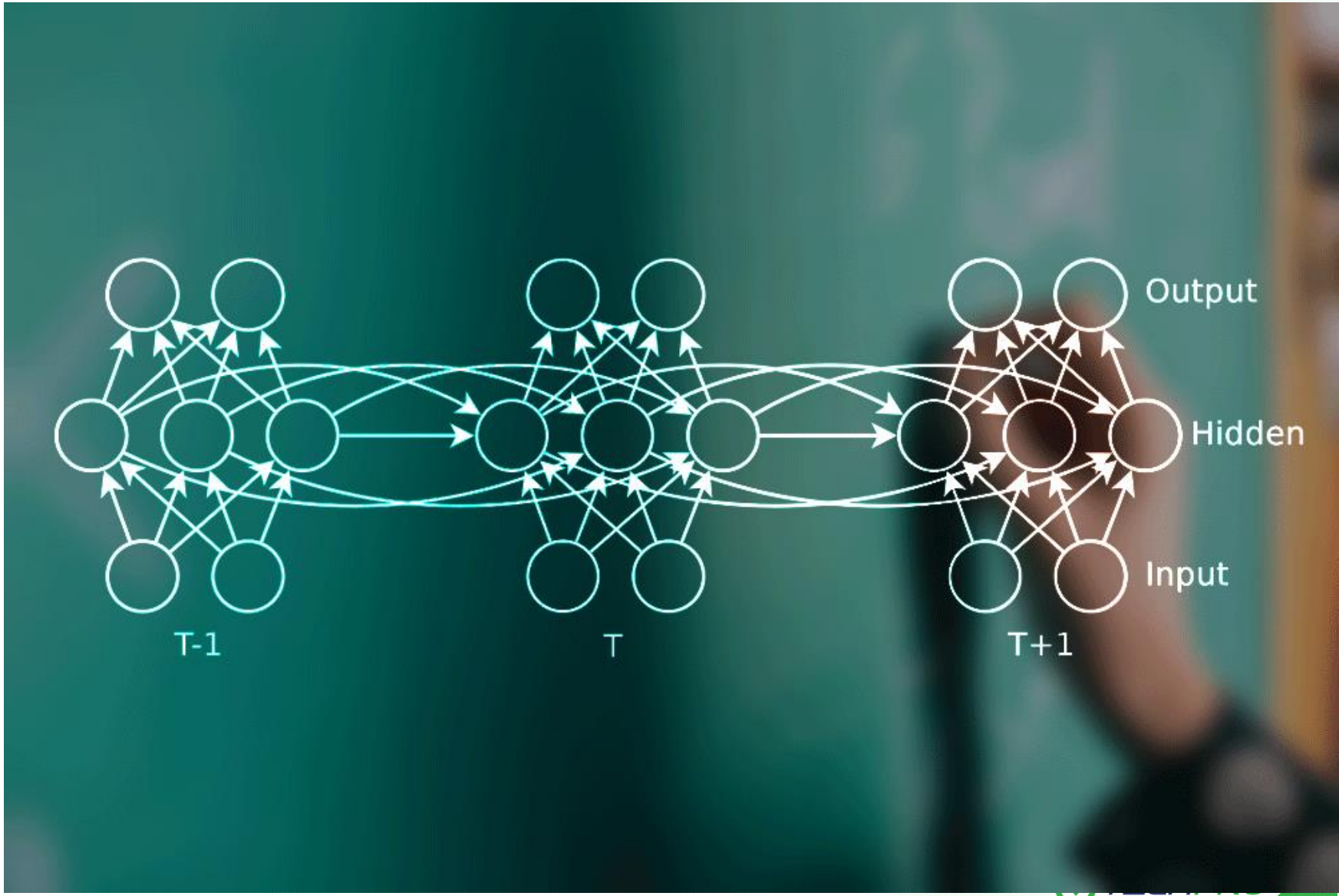
Machine Translation

Given an input in one language, RNN can be used to translate the input into different languages as output

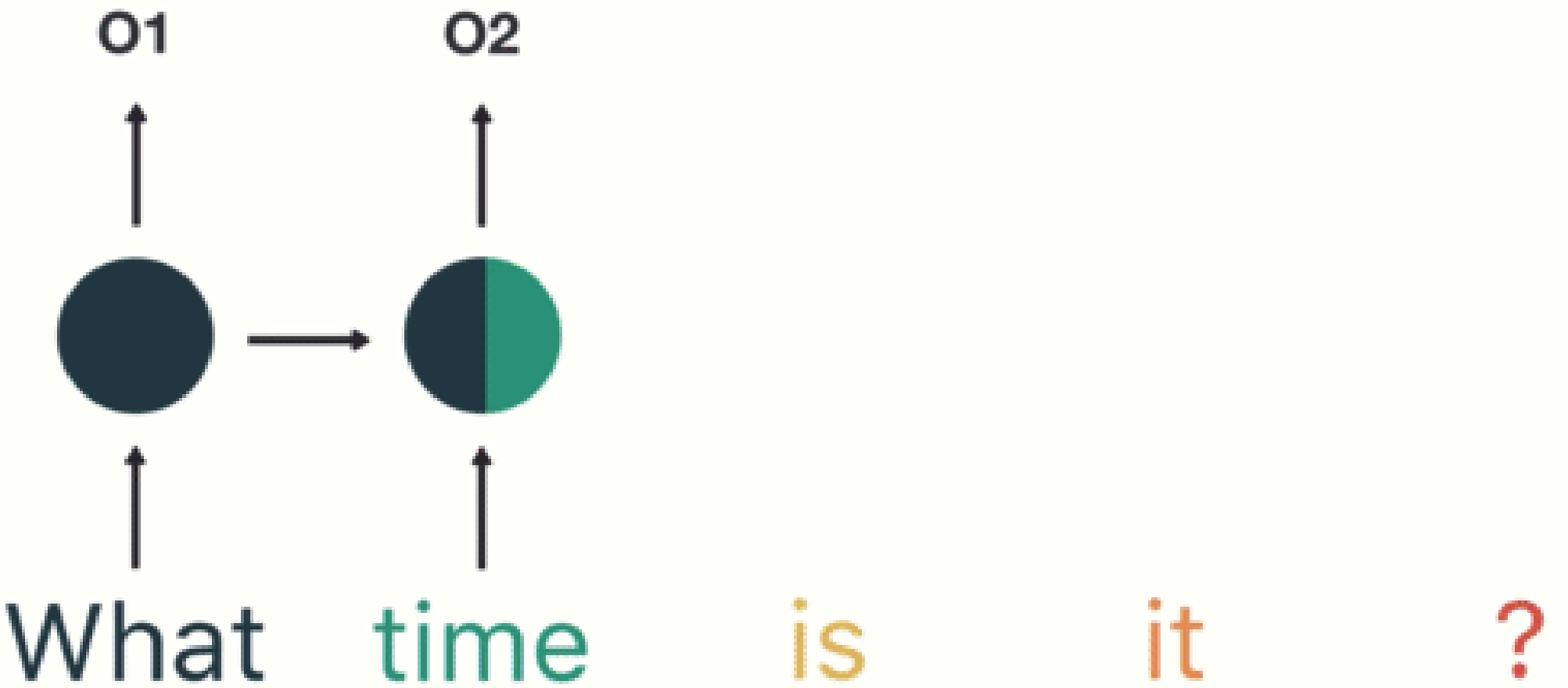
HOW DOES RNN WORK?

RNN

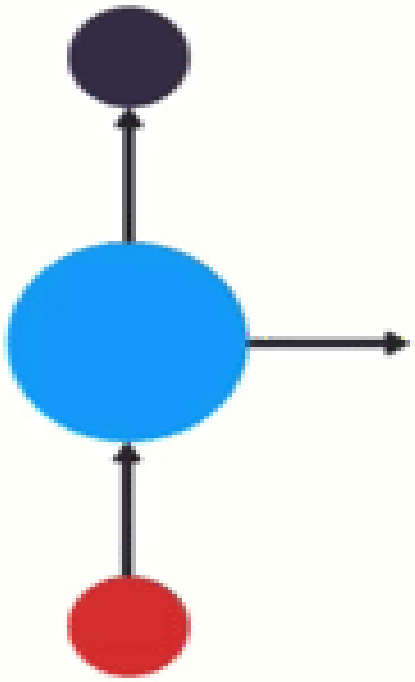
A **recurrent neural network** (**RNN**) is a type of artificial neural network which uses sequential data or time series data.



RNN



RNN



Machine Learning

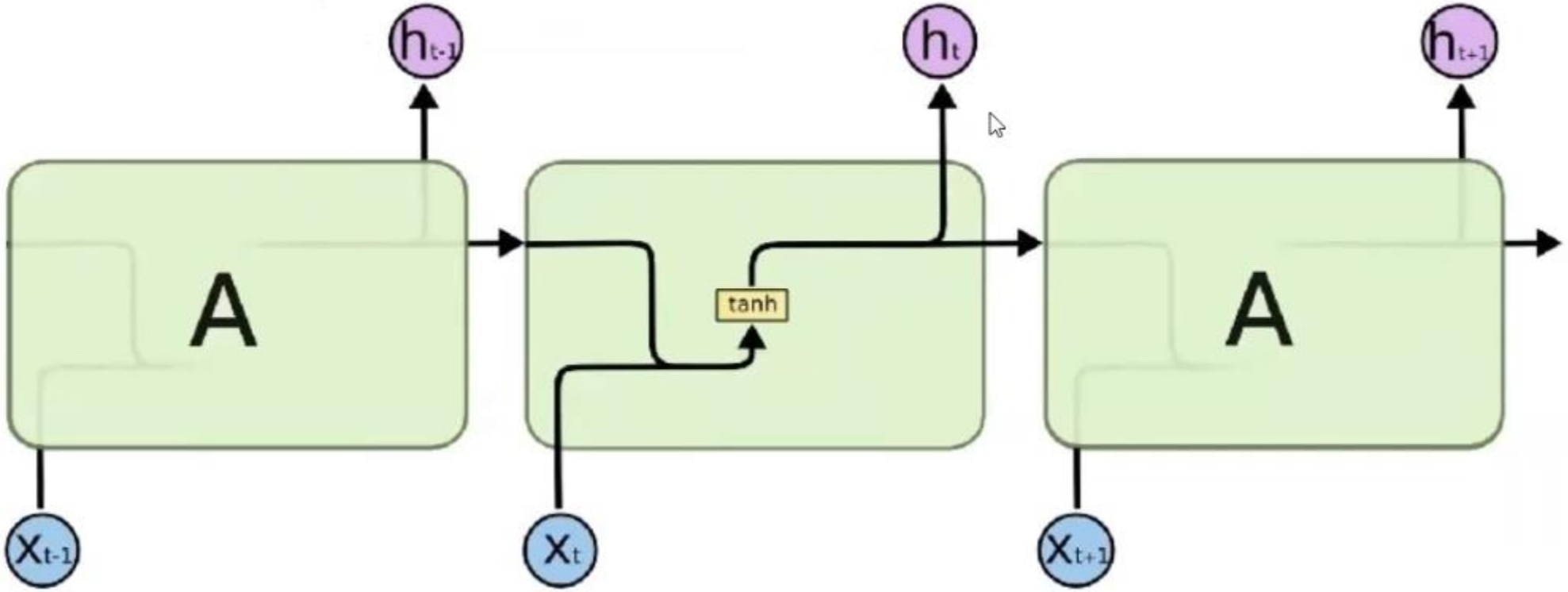
time_series_generator

data = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

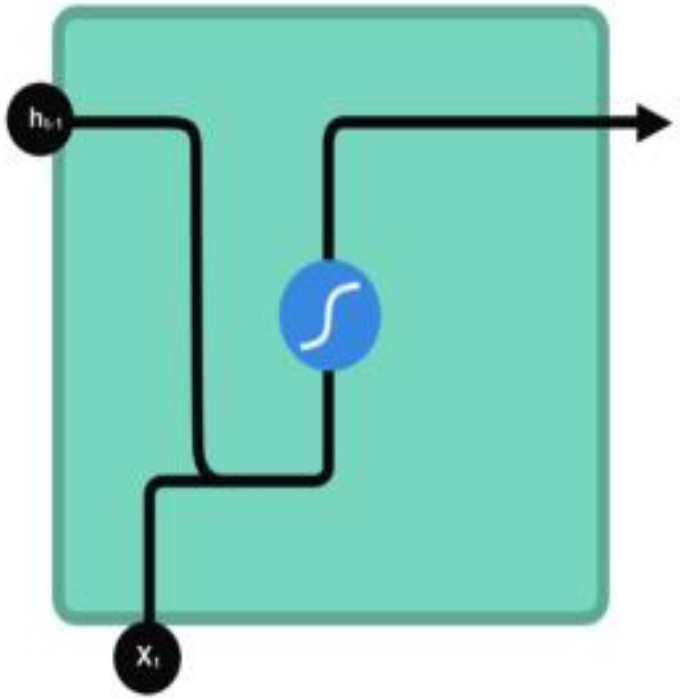
```
Grup 1 - Girdi (x): [[1 2 3]], Çıkış (y): [4]
Grup 2 - Girdi (x): [[2 3 4]], Çıkış (y): [5]
Grup 3 - Girdi (x): [[3 4 5]], Çıkış (y): [6]
Grup 4 - Girdi (x): [[4 5 6]], Çıkış (y): [7]
Grup 5 - Girdi (x): [[5 6 7]], Çıkış (y): [8]
Grup 6 - Girdi (x): [[6 7 8]], Çıkış (y): [9]
Grup 7 - Girdi (x): [[7 8 9]], Çıkış (y): [10]
```


```
batch_size = 1
```

RNN

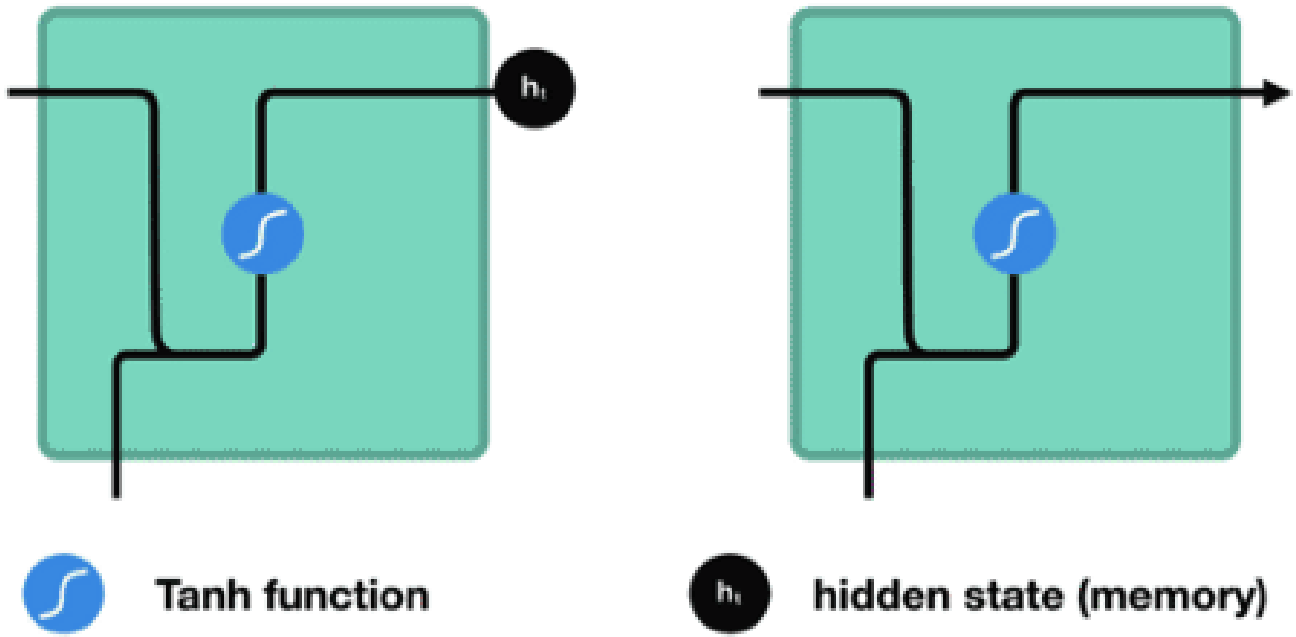


RNN

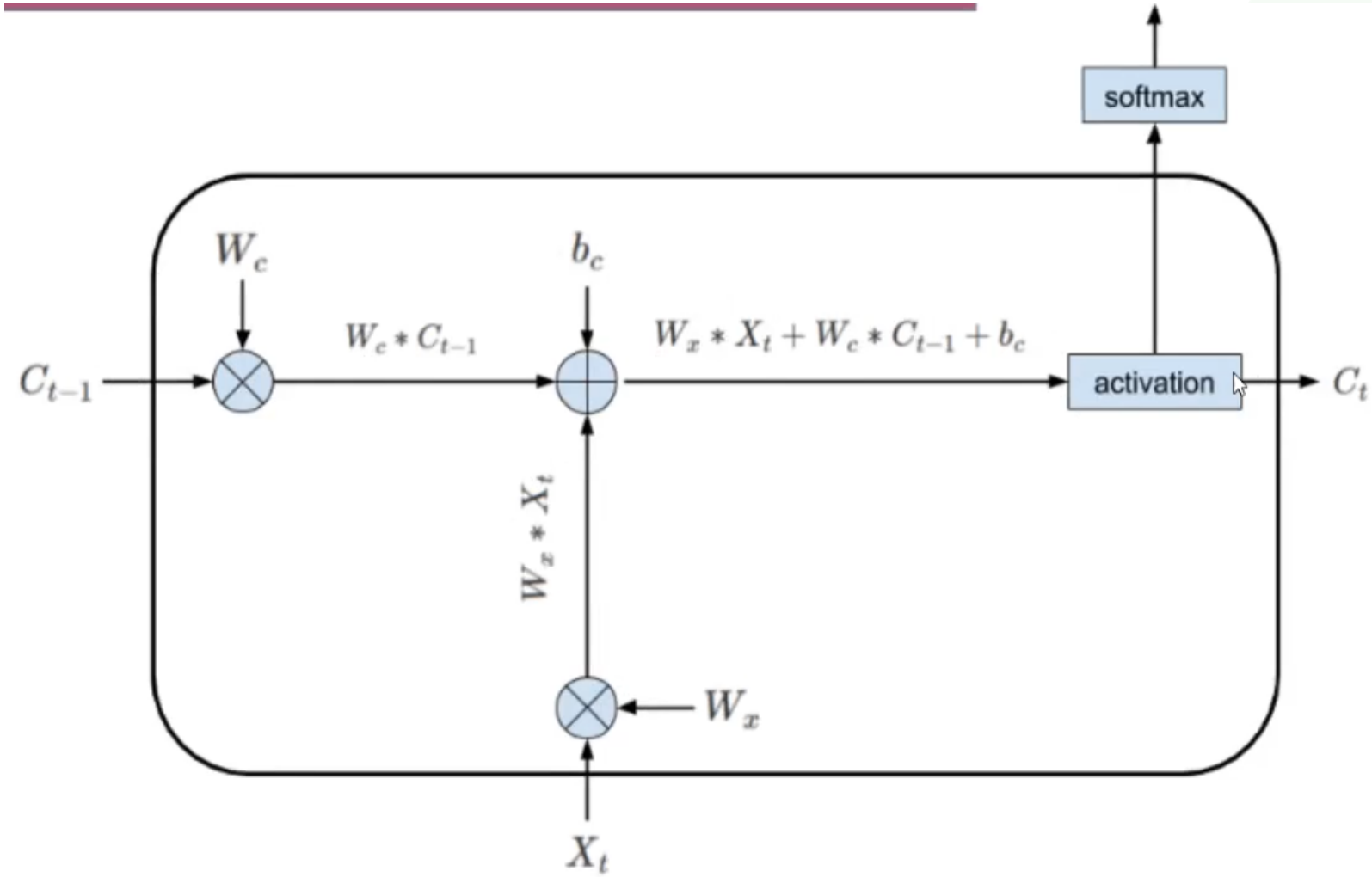


- Tanh function
- h_t new hidden state
- h_{t-1} previous hidden state
- x_t input
- concatenation

RNN



RNN



X_t	Is the input to the RNN at the time t
C_{t-1}	Is the cell state at time t-1
W_c	Is the weight matrix for C
b_c	Is the bias matrix
W_x	Is the weight matrix for X
C_t	Is the cell state at time t
Y_t	Is the output (effectively the prediction) of the RNN
⊗	Represents matrix multiplication
⊕	Represents matrix addition
activation	Is the activation function (e.g. tanh activation would constrain values between 1 & -1)

DATE : 15.10.2024

DT/NT : DT

LESSON : DEEP LEARNING

SUBJECT: RNN– LONG SHORT-TERM MEMORY(LSTM) - GRU

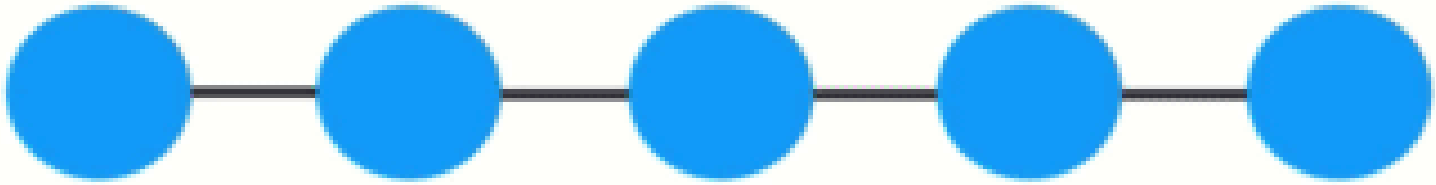
BATCH : 250

DATA SCIENCE

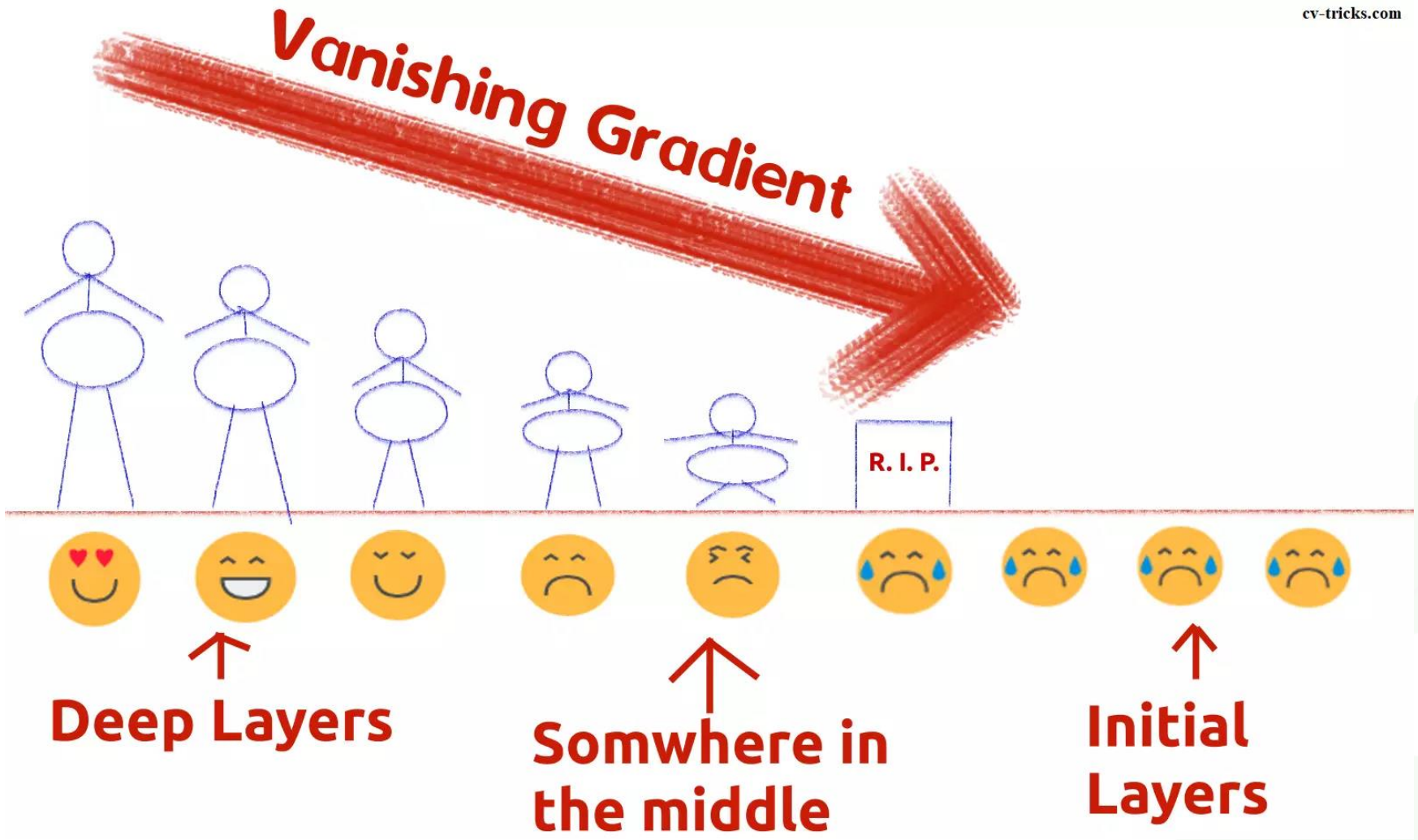
TECHPRO EDUCATION

VANISHING /EXPLODING GRADIENT PROBLEM

RNN



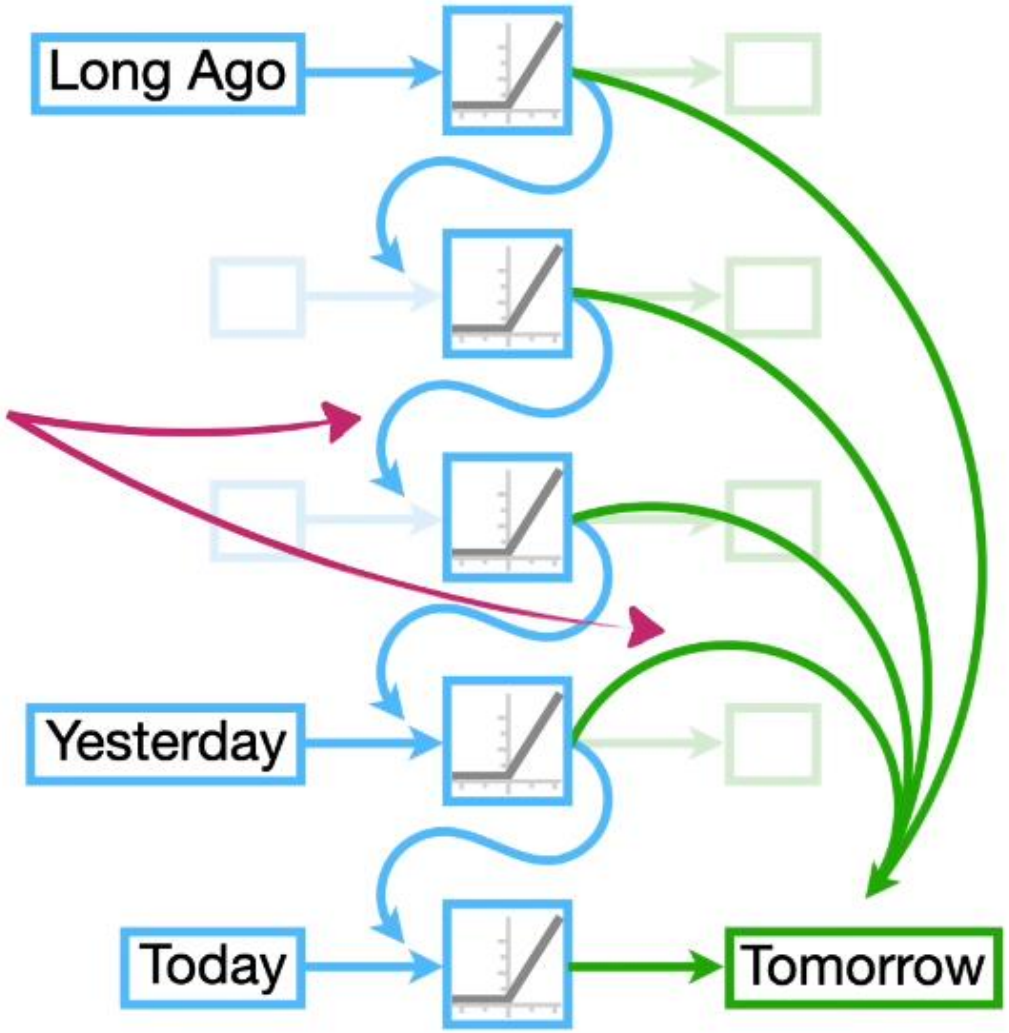
RNN VANISHING /EXPLODING GRADIENT



LSTM (LONG SHORT-TERM MEMORY)

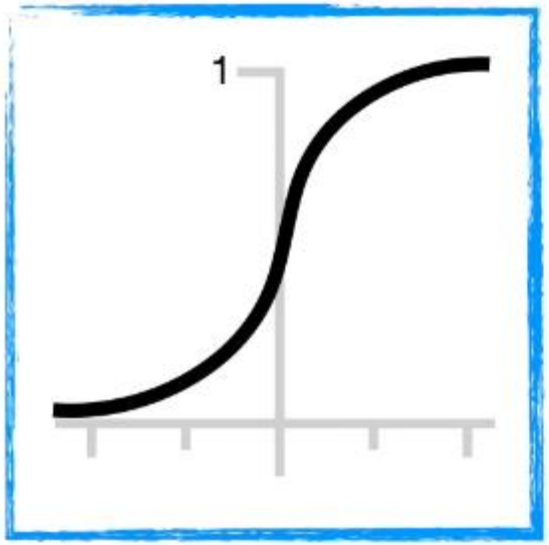
LSTM

…**Long Short-Term Memory** uses two separate paths to make predictions about tomorrow.

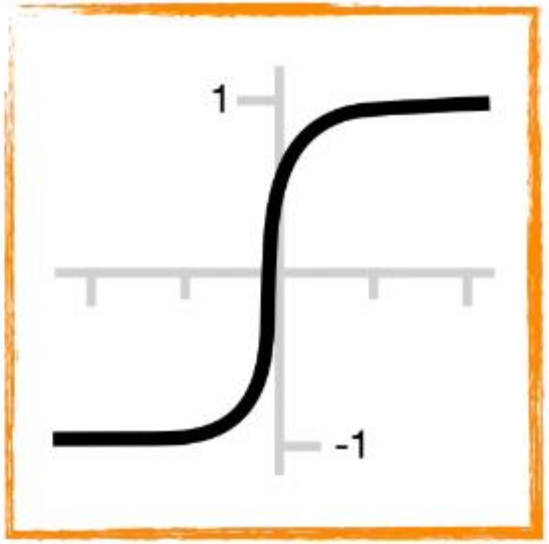


LSTM

So, now that we know that the **Sigmoid Activation Function** turns any input into a number between **0** and **1**...



...and the **Tanh Activation Function** turns any input into a number between **-1** and **1**...



LSTM

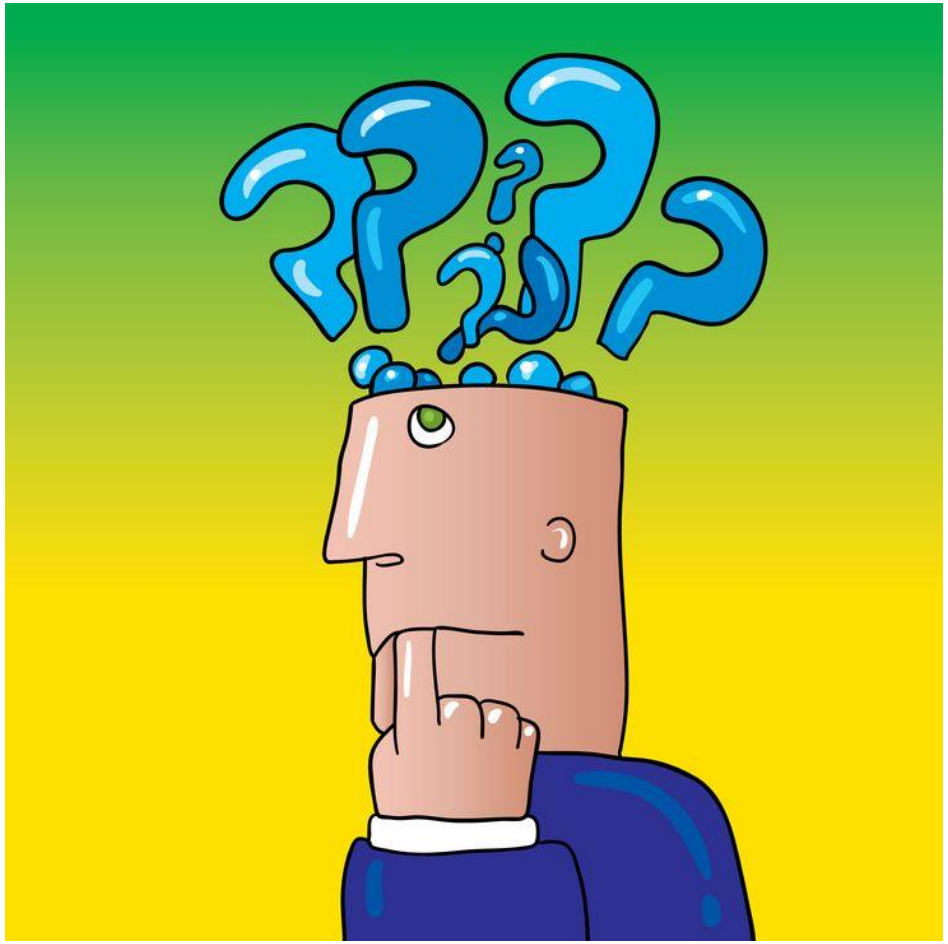
Machine Learning

What is LSTM?

TECHPRO EDUCATION

LSTM

"I will lead the way to success.."
“She wore a necklace made of lead."



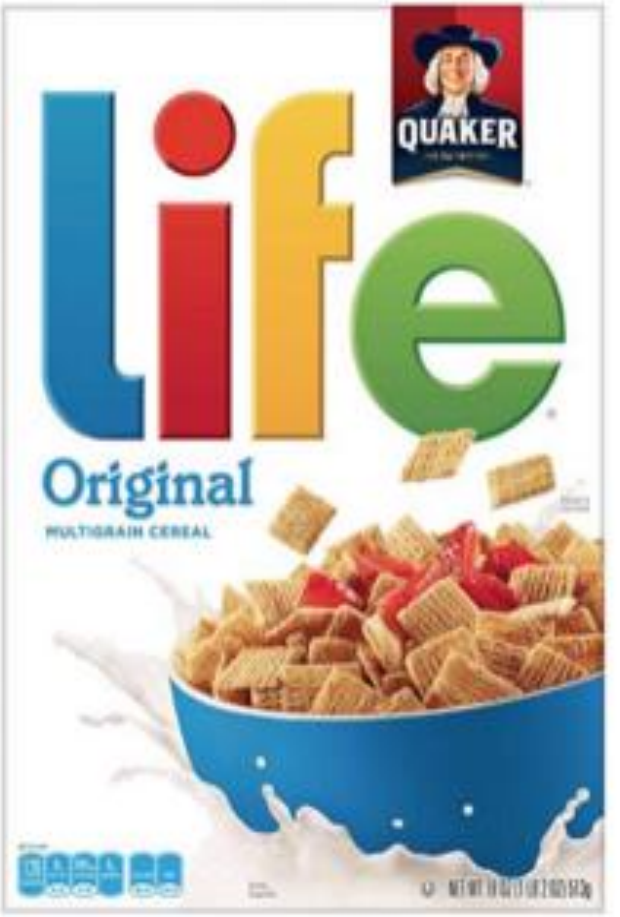
Customers Review 2,491

Thanos

September 2018

Verified Purchase

Amazing! This box of cereal gave me a perfectly balanced breakfast, as all things should be. I only ate half of it but will definitely be buying again!



A Box of Cereal

$3.99

Customers Review 2,491

Thanos

September 2018

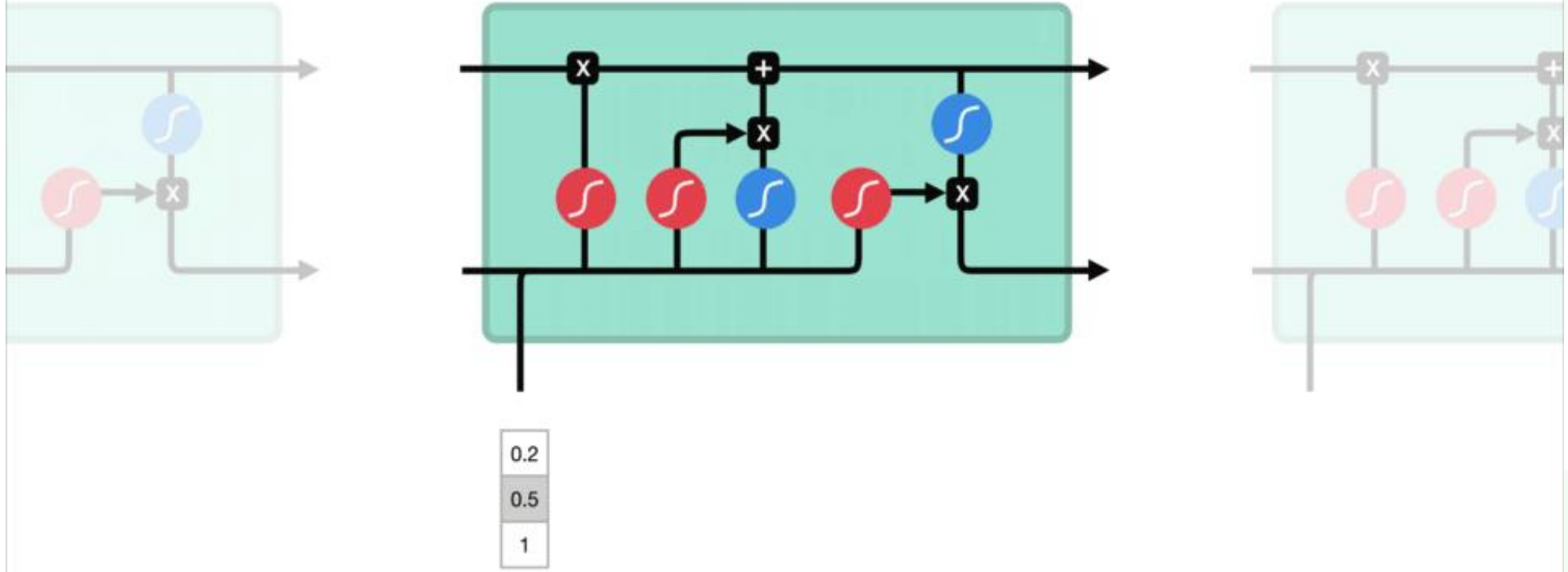
Verified Purchase

Amazing! This box of cereal gave me a perfectly balanced breakfast, as all things should be. I only ate half of it but will definitely be buying again!

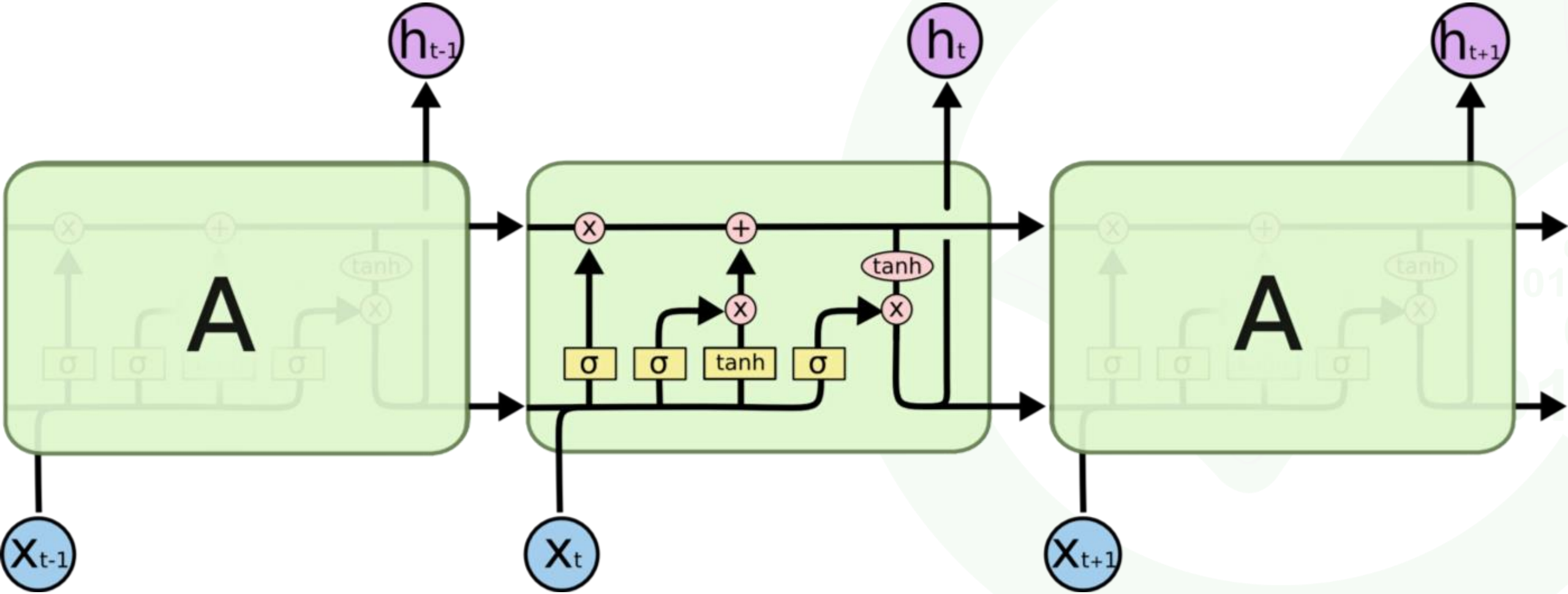
A Box of Cereal

$3.99

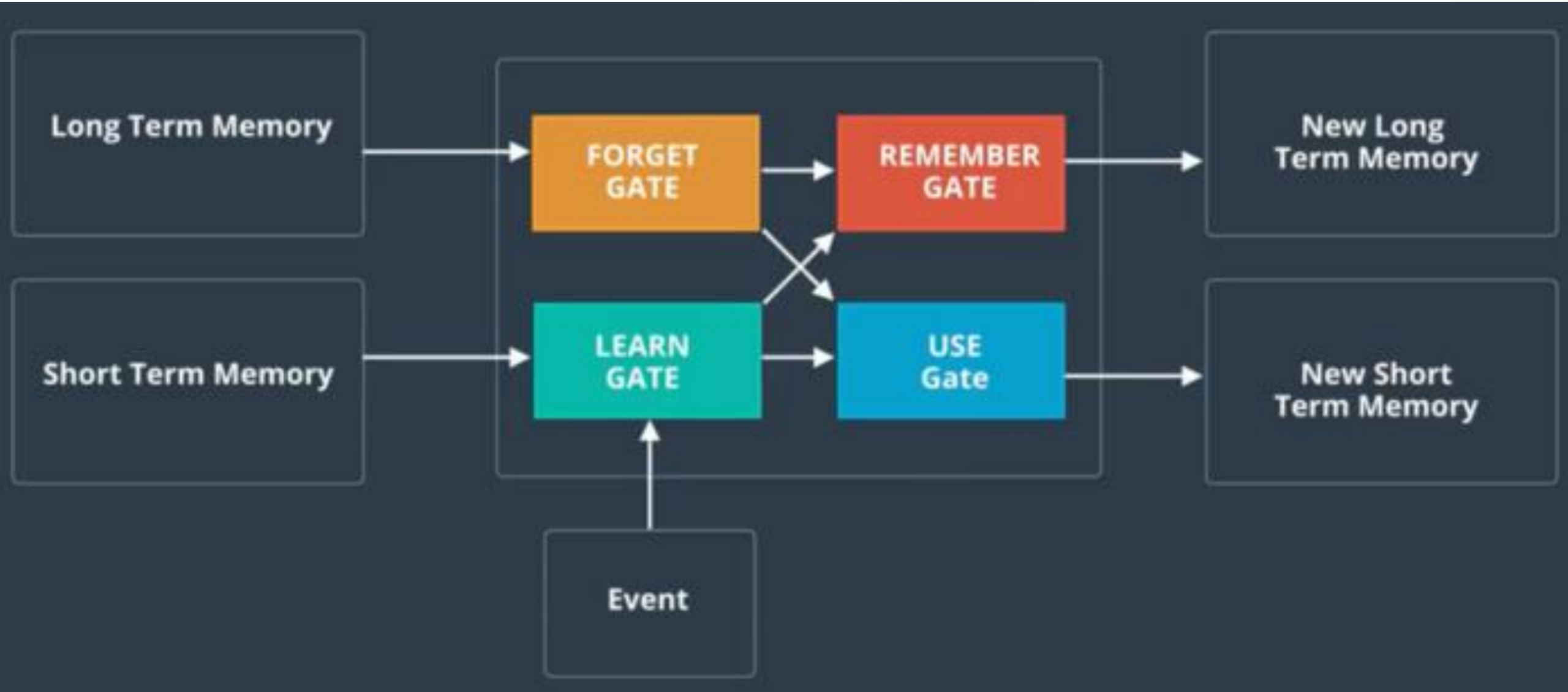
LSTM



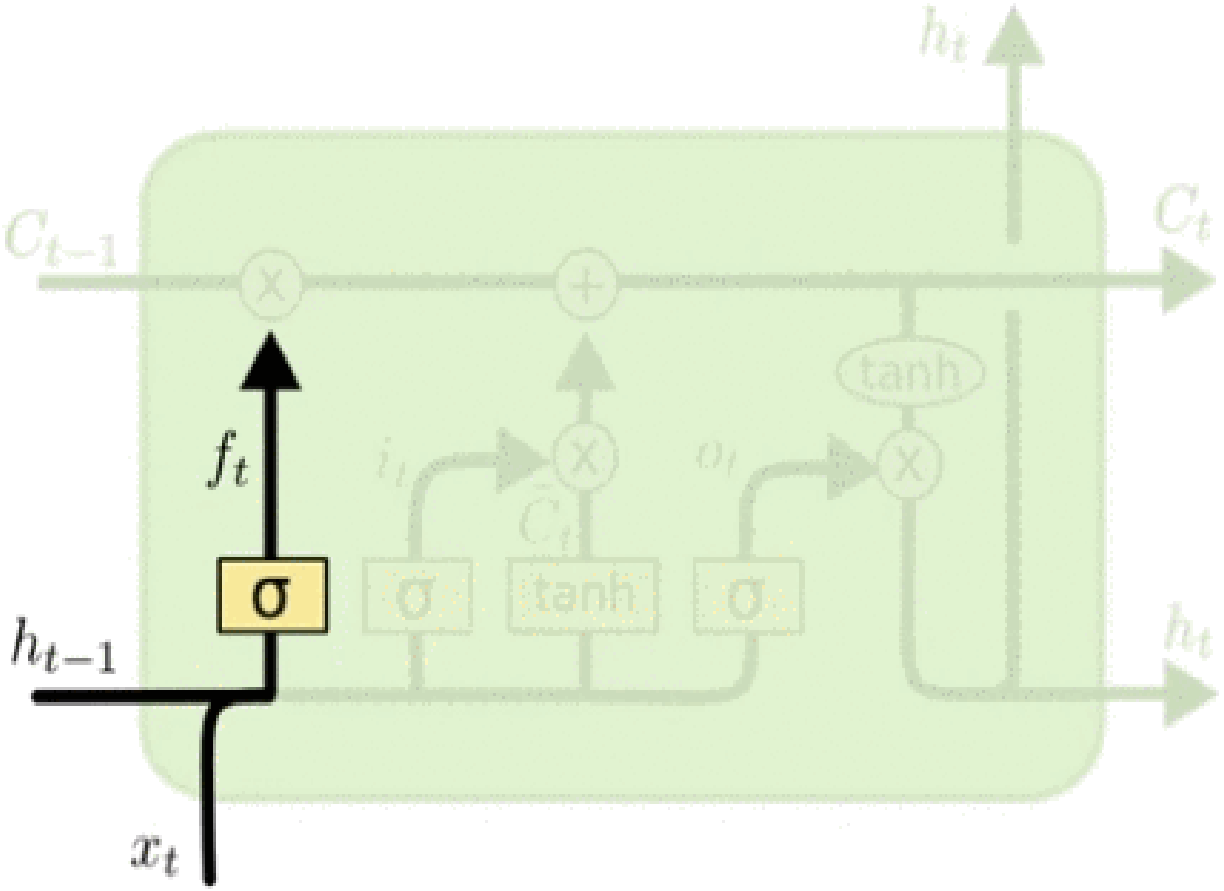
LSTM



LSTM

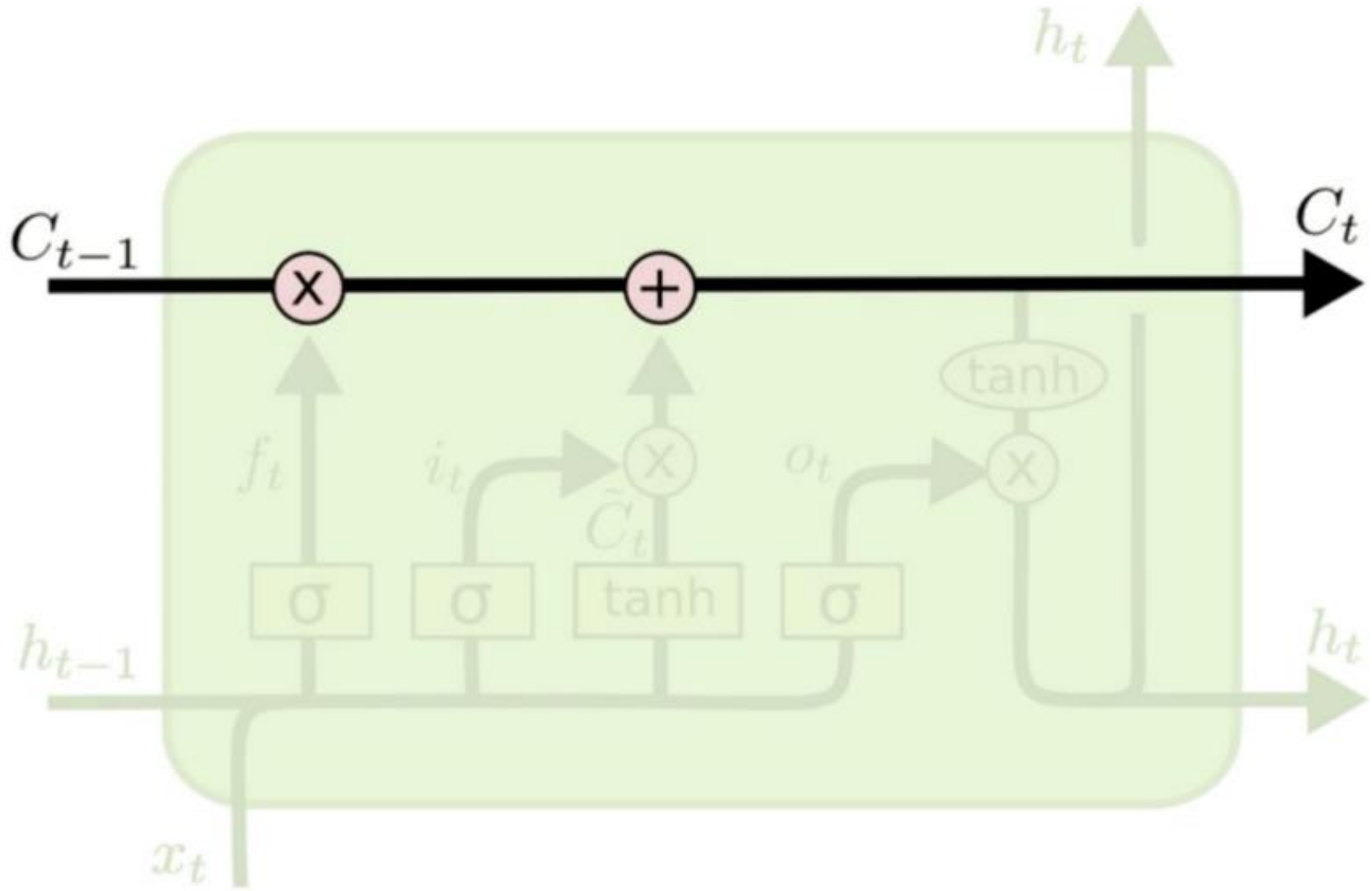

Long Term Memory
Short Term Memory
FORGET GATE
REMEMBER GATE
LEARN GATE
USE Gate
Event
New Long Term Memory
New Short Term Memory

LSTM



(1) $$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

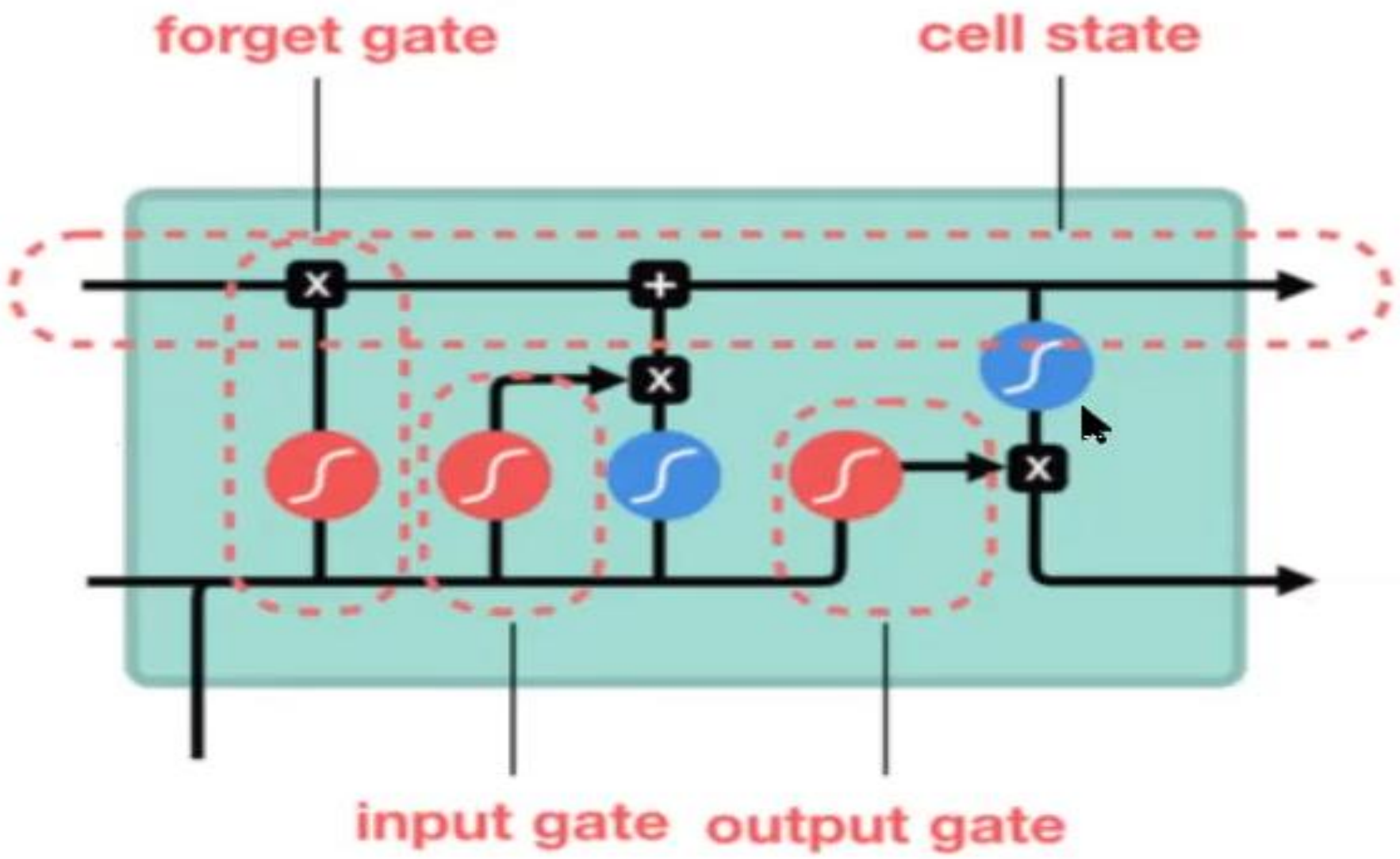
LSTM



LSTM

sigmoid

tanh

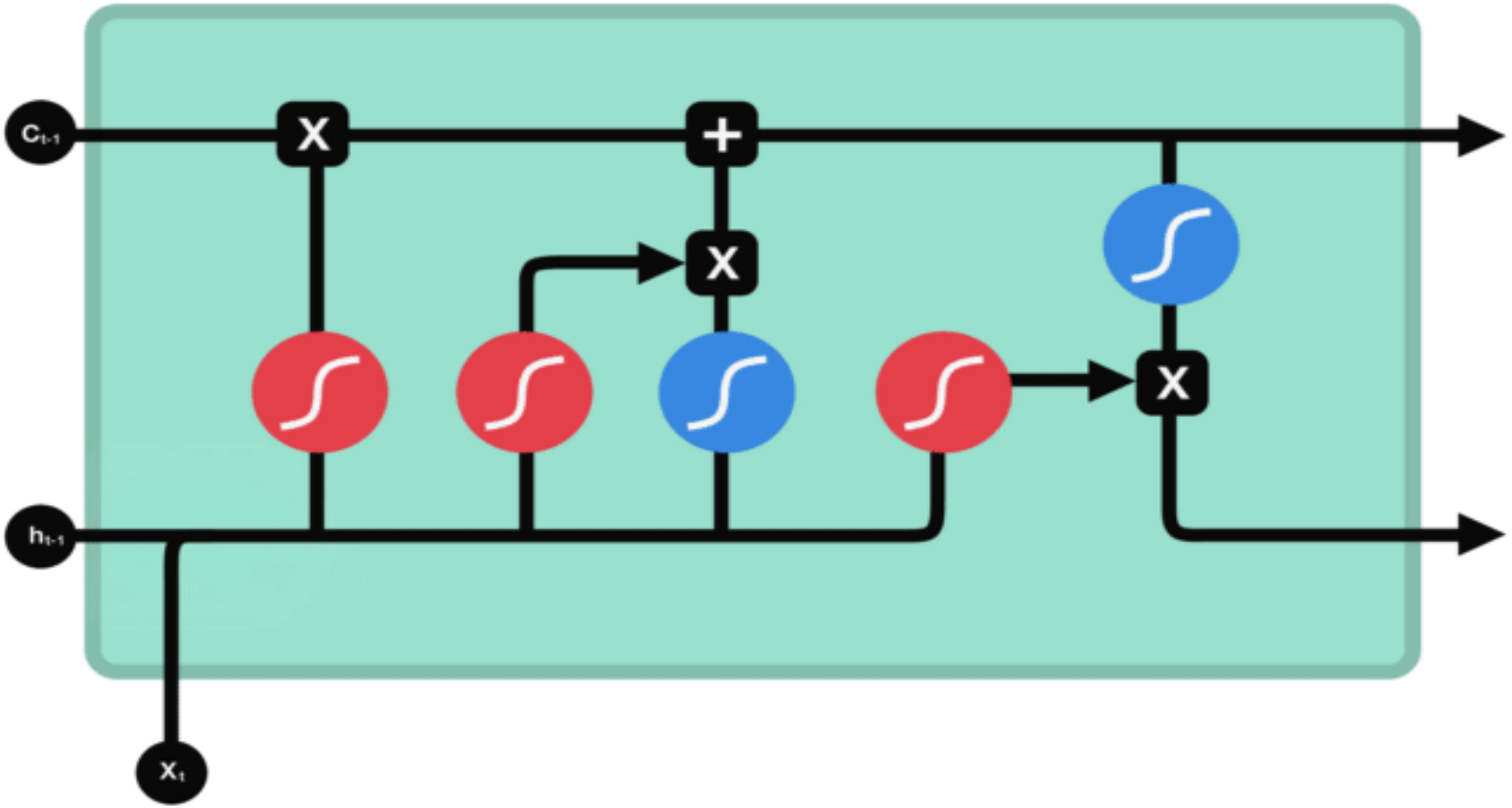


LSTM

1-) Forget Gate

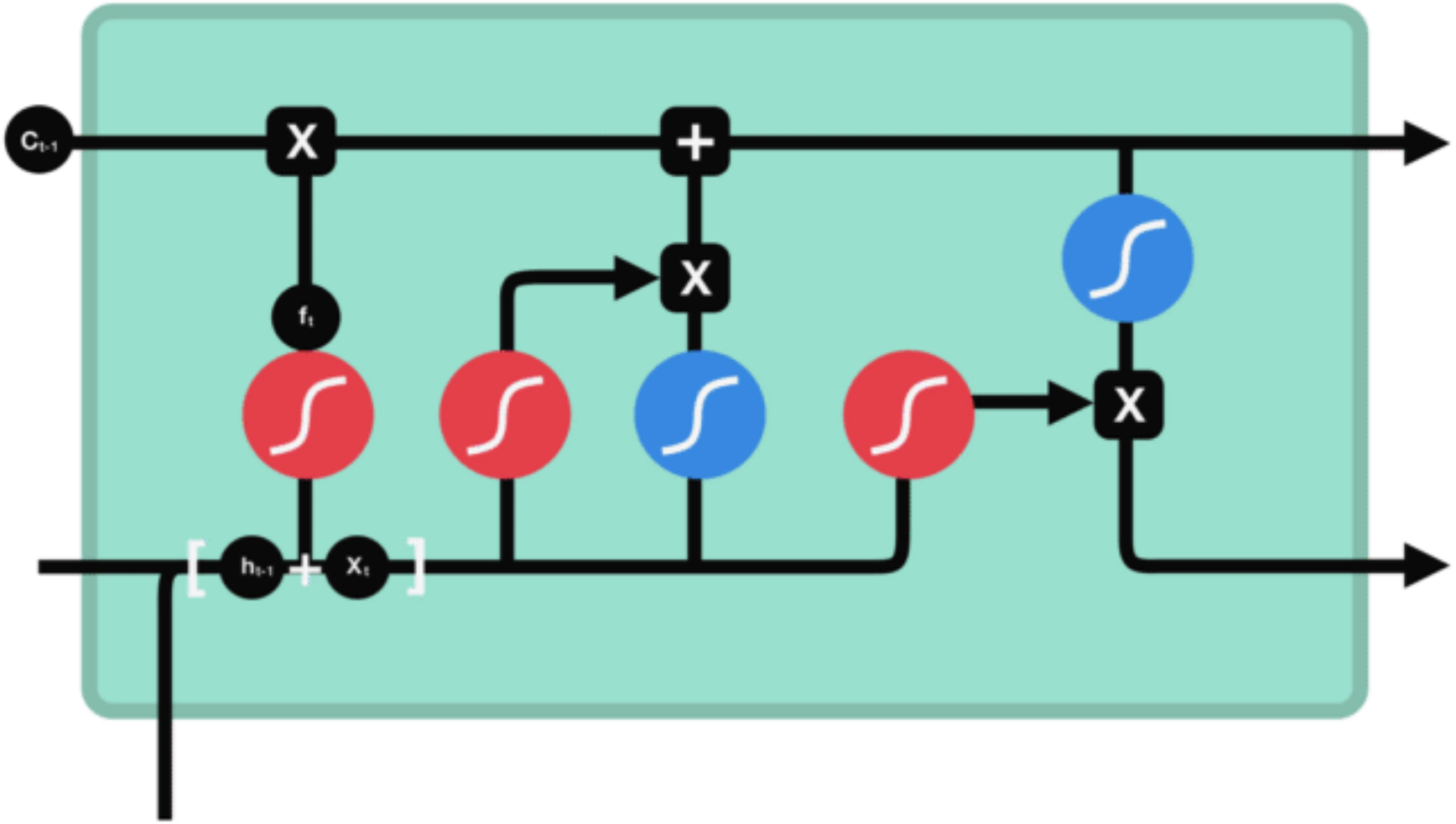
c_{t-1} previous cell state

f_t forget gate output



LSTM

2-) Input Gate



C_{t-1} previous cell state

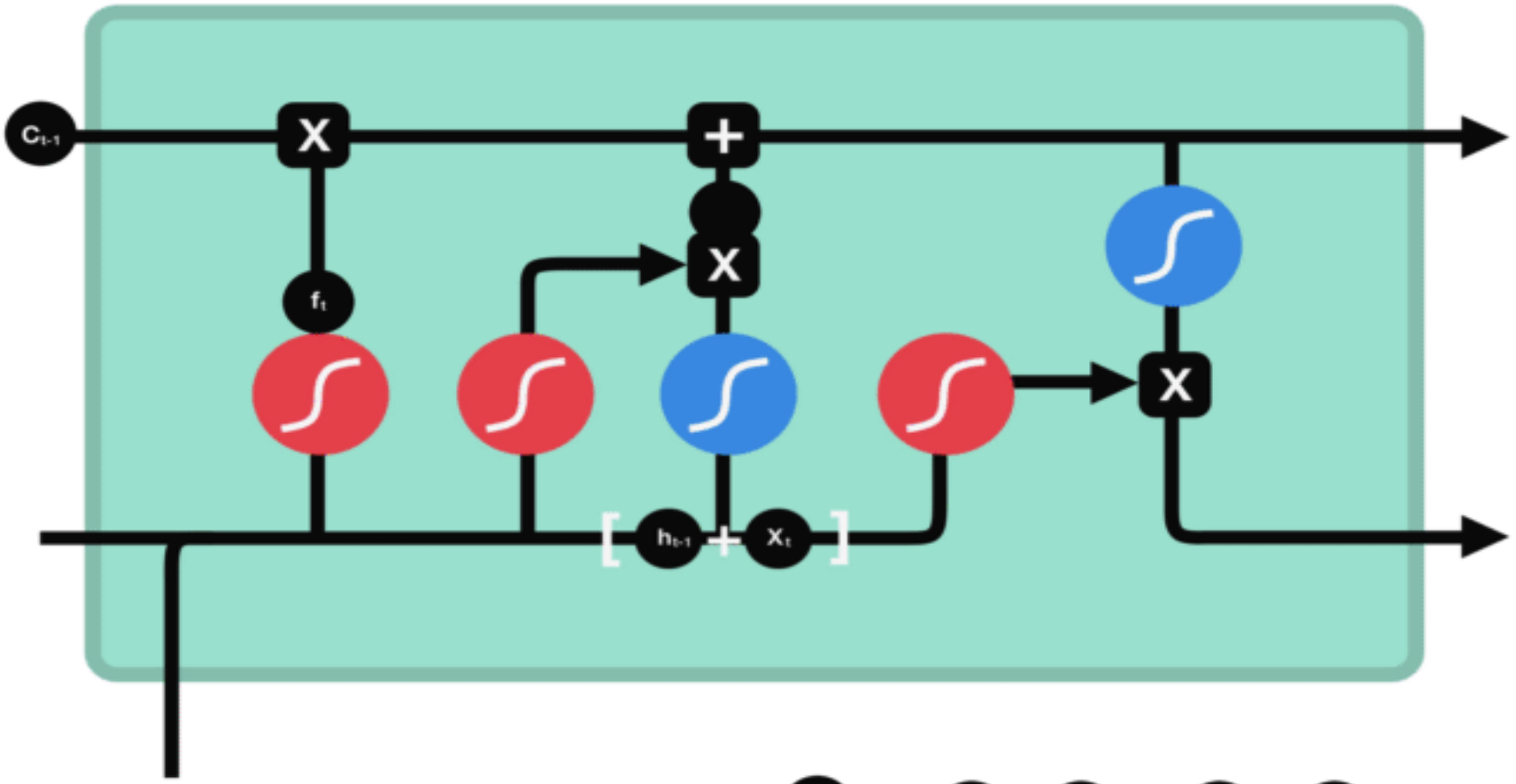
f_t forget gate output

i_t input gate output

$\tilde{C}_t$ candidate

LSTM

3-) Cell State

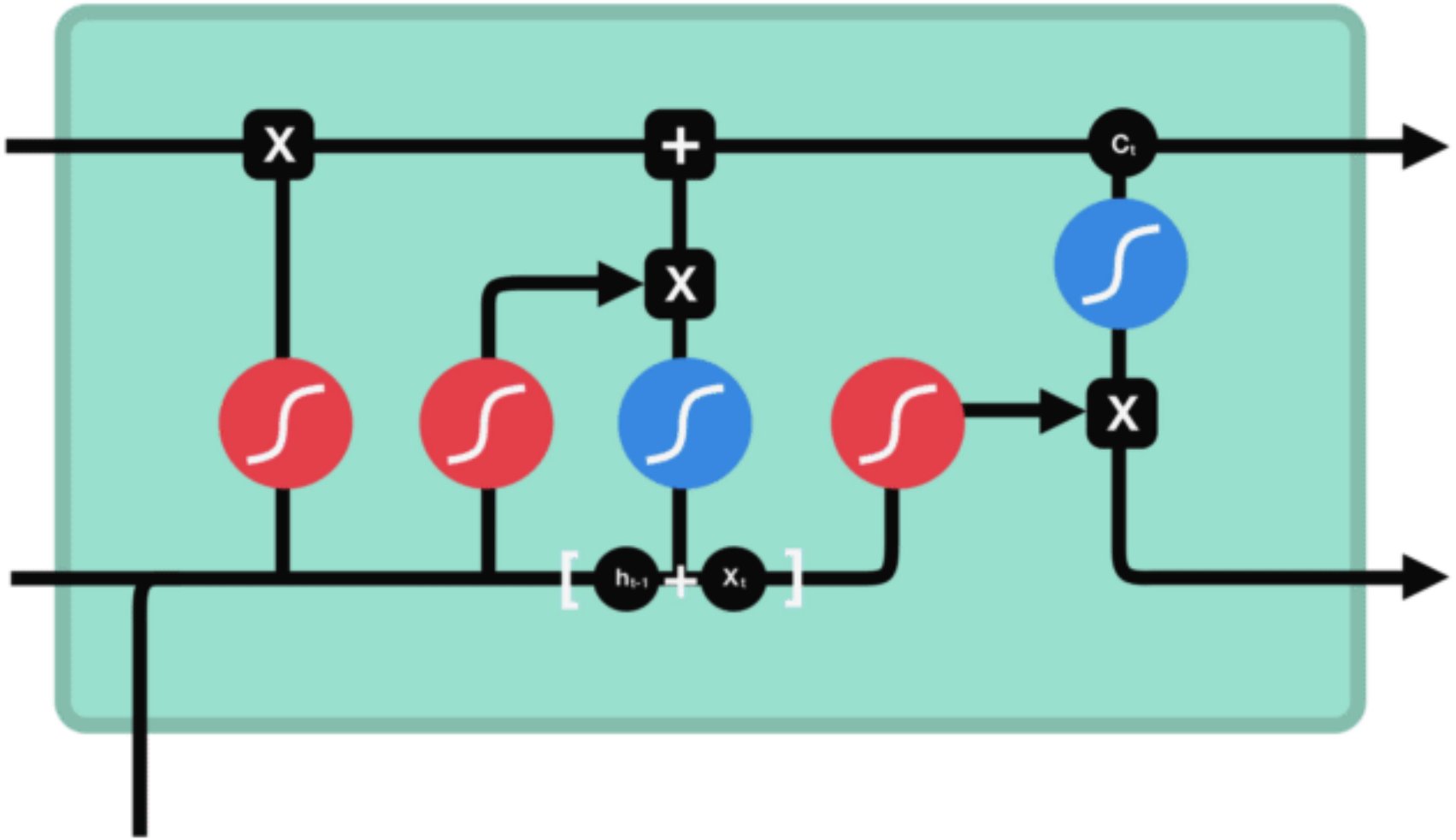


- C_{t-1} previous cell state
- f_t forget gate output
- i_t input gate output
- $\tilde{C}_t$ candidate
- C_t new cell state

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

LSTM

4-) Output Gate



c_{t-1} previous cell state

f_t forget gate output

i_t input gate output

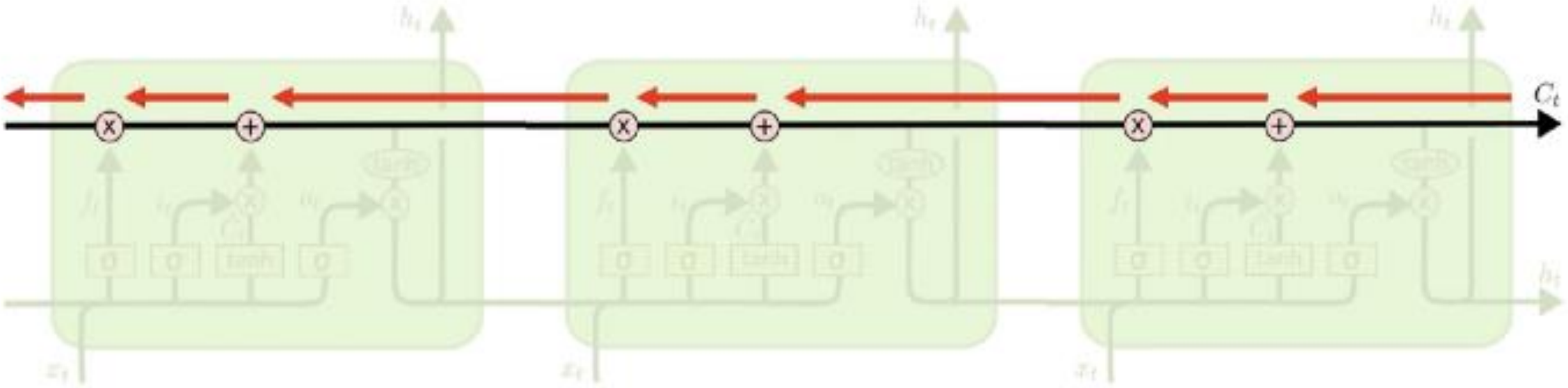
$\tilde{c}_t$ candidate

c_t new cell state

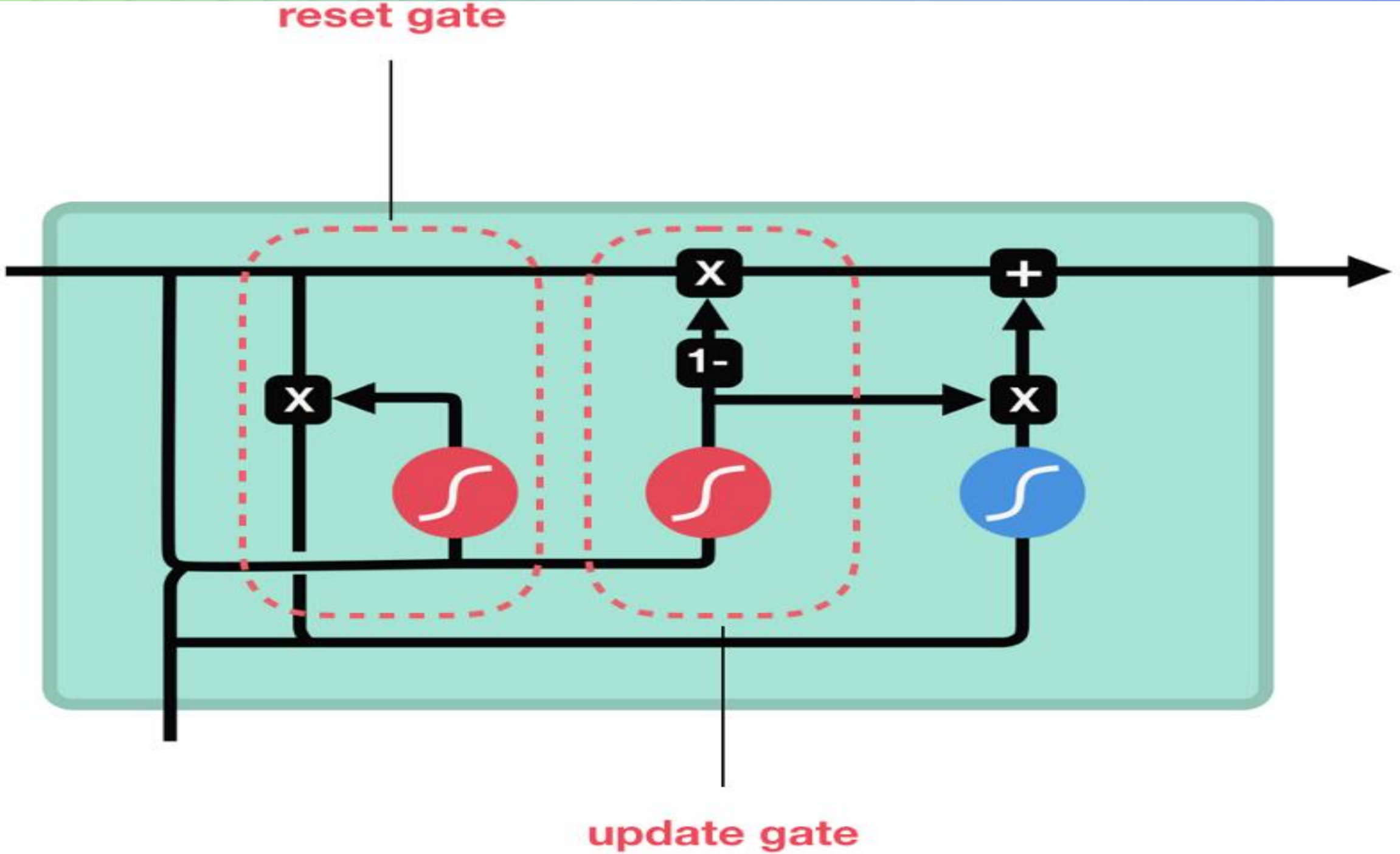
o_t output gate output

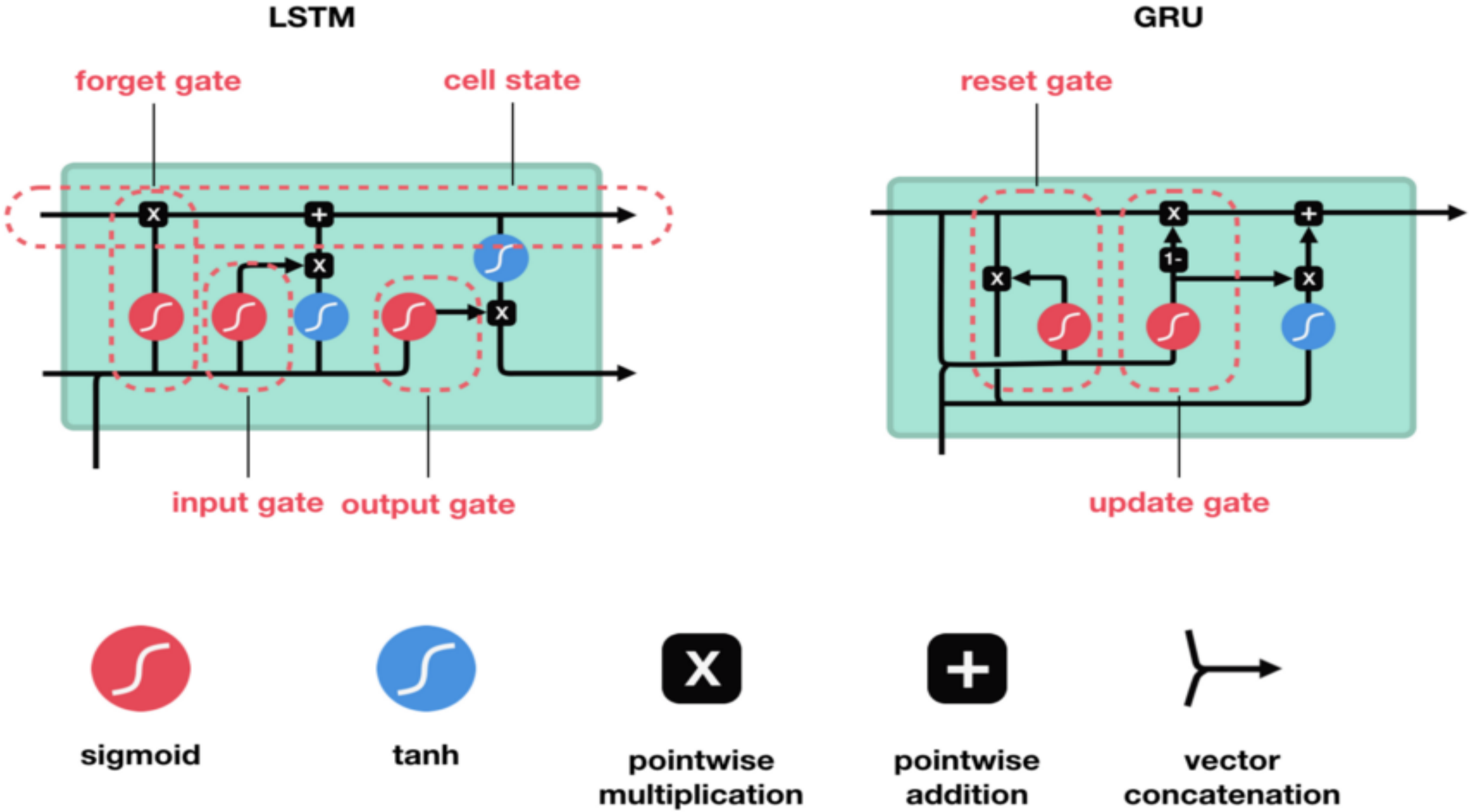
h_t hidden state

LSTM - BACKPROPAGATION

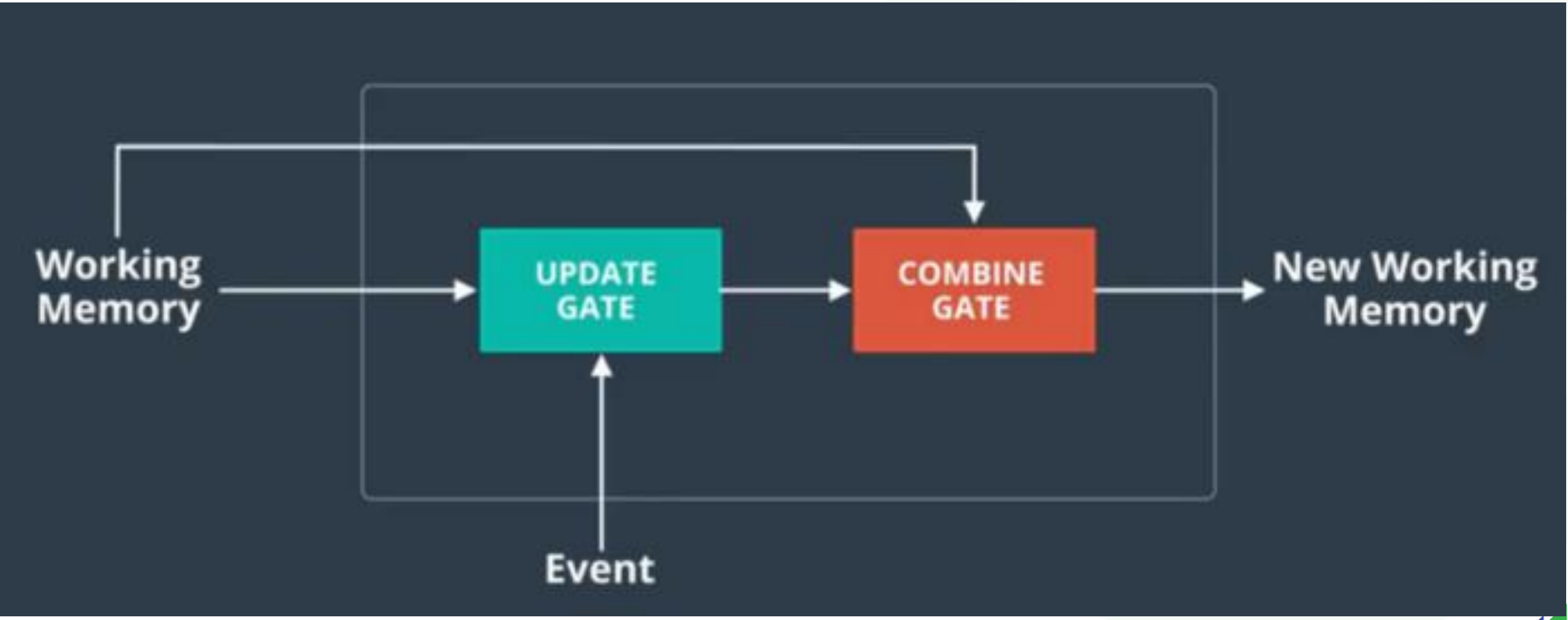


GRU (GATED RECCURENT UNITS)


reset gate
X
+
1-
X
X
update gate

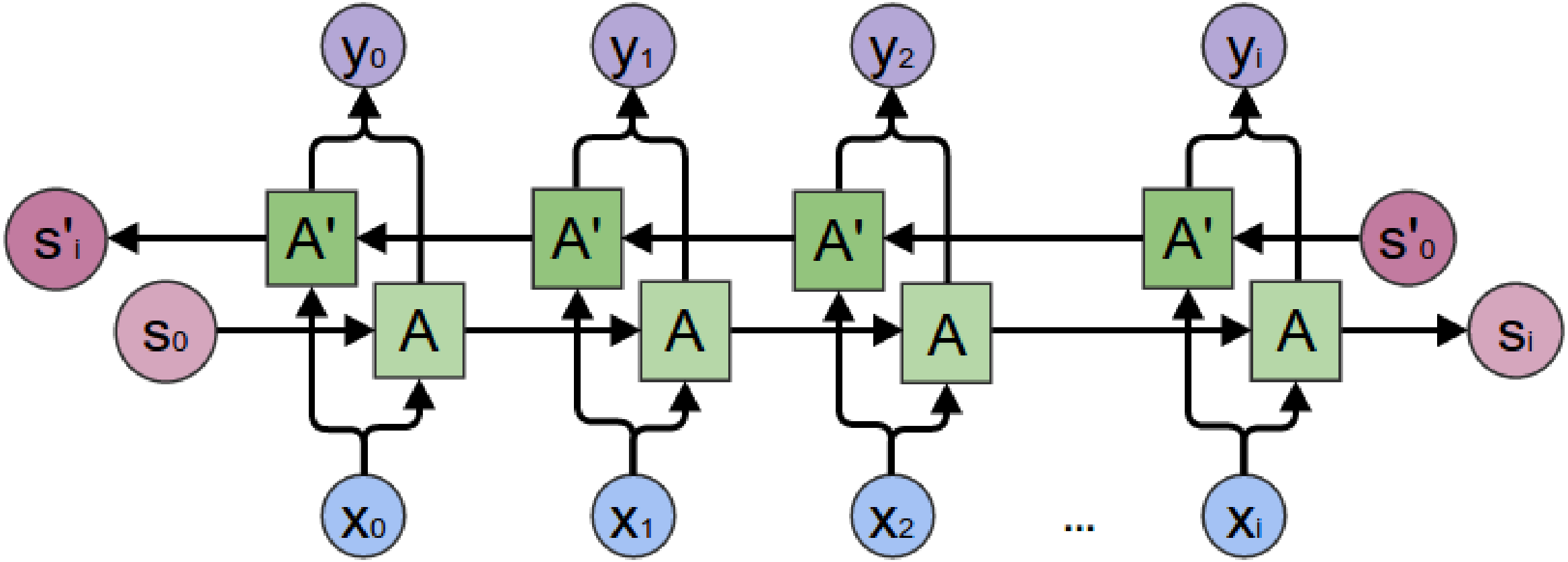

LSTM
forget gate
cell state
input gate
output gate
GRU
reset gate
update gate
1-
sigmoid
tanh
pointwise multiplication
pointwise addition
vector concatenation

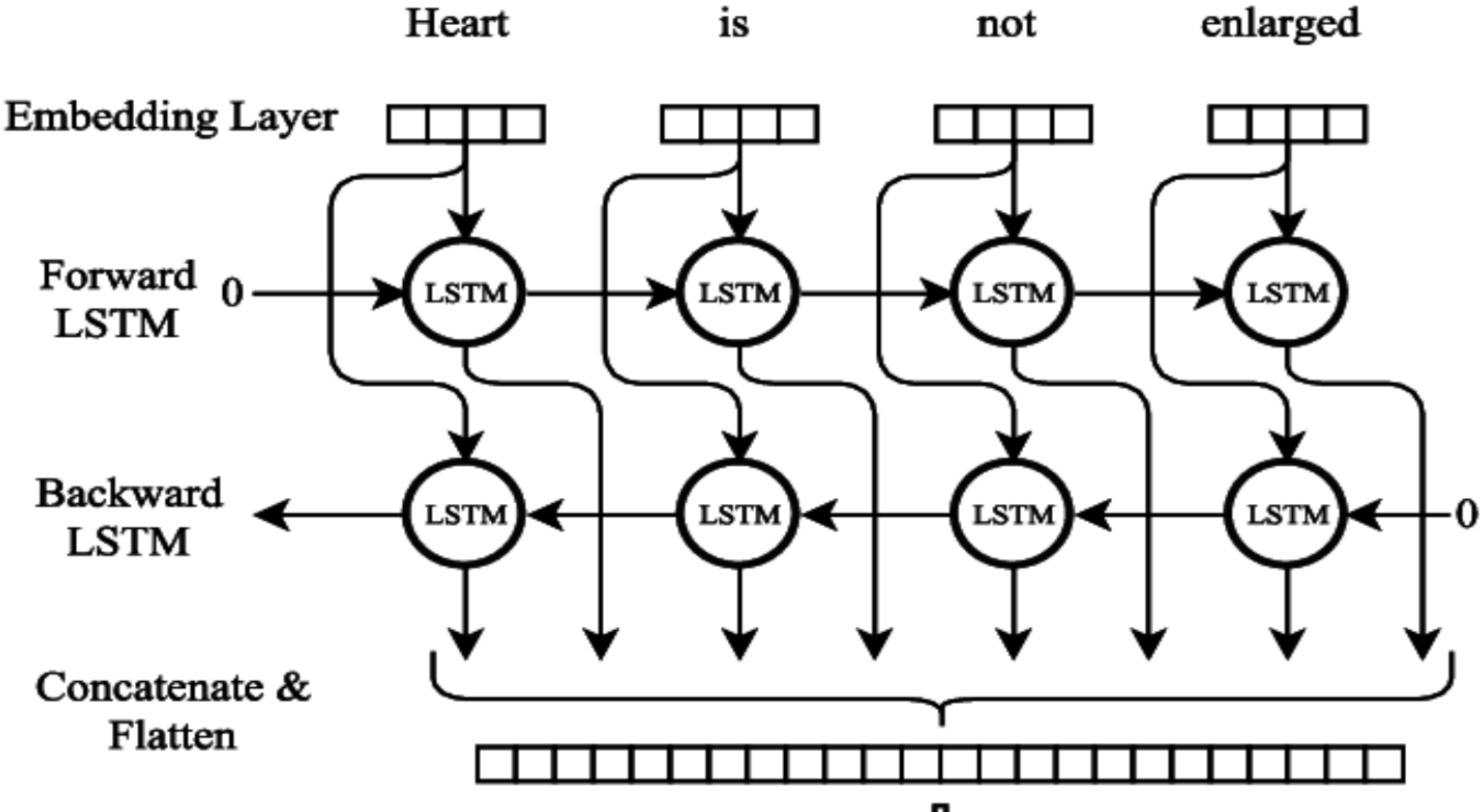
GRU


Working Memory
UPDATE GATE
COMBINE GATE
New Working Memory
Event

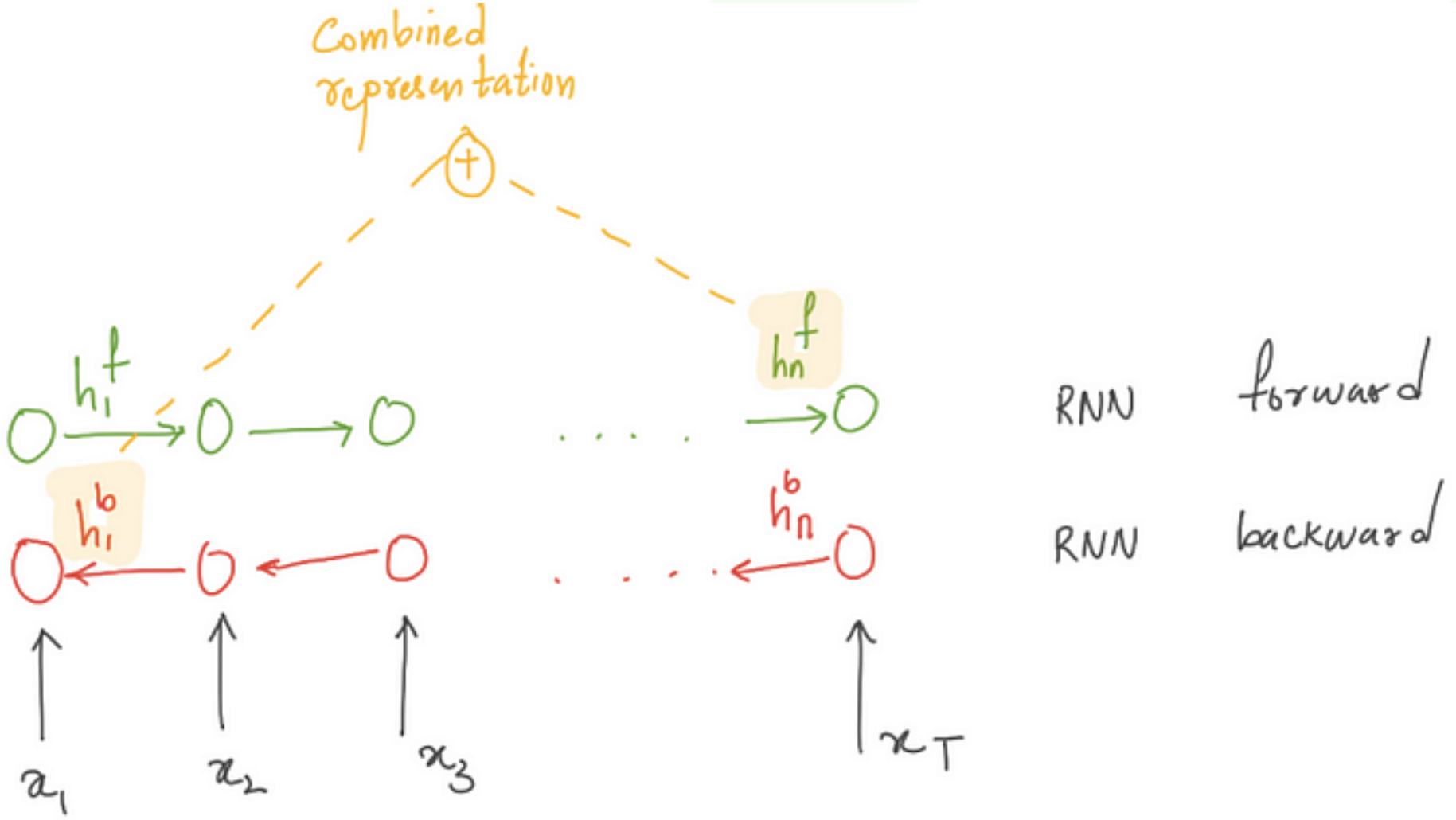
BIDIRECTIONAL RNN/LSTM/GRU

BIDIRECTIONAL RNN/LSTM/GRU

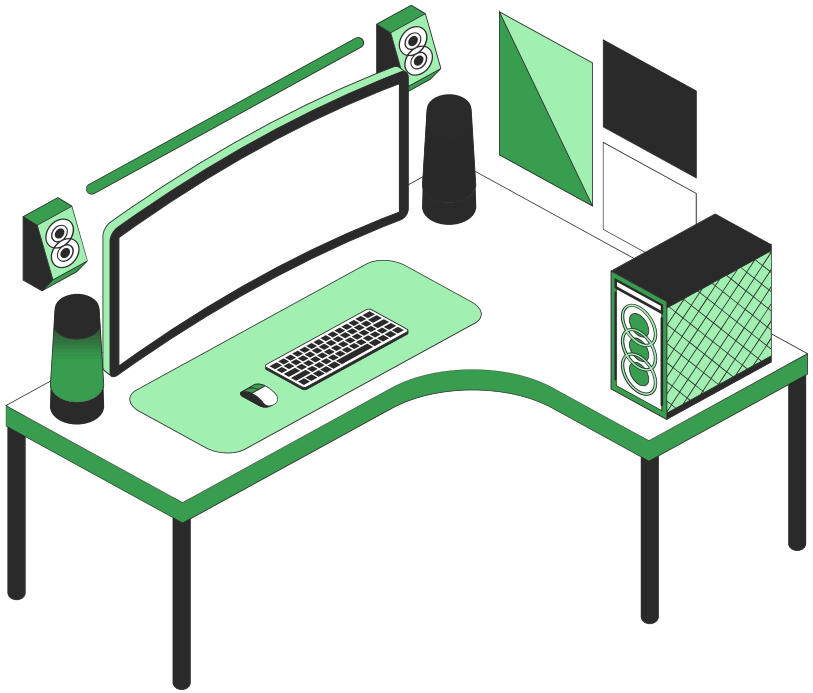



Heart
is
not
enlarged
Embedding Layer
Forward LSTM
0
LSTM
LSTM
LSTM
LSTM
Backward LSTM
LSTM
LSTM
LSTM
LSTM
0
Concatenate & Flatten

BIDIRECTIONAL RNN/LSTM/GRU



THE END



Do you have any questions?

Send it to us! We hope you learned something new.

Tea break...

00:00