Apple Stock Prediction

Soheil Tehranipour

August 2023

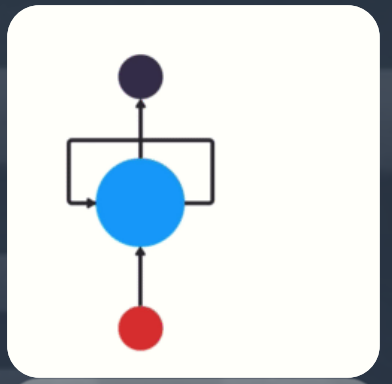
سهیل تهرانی پور

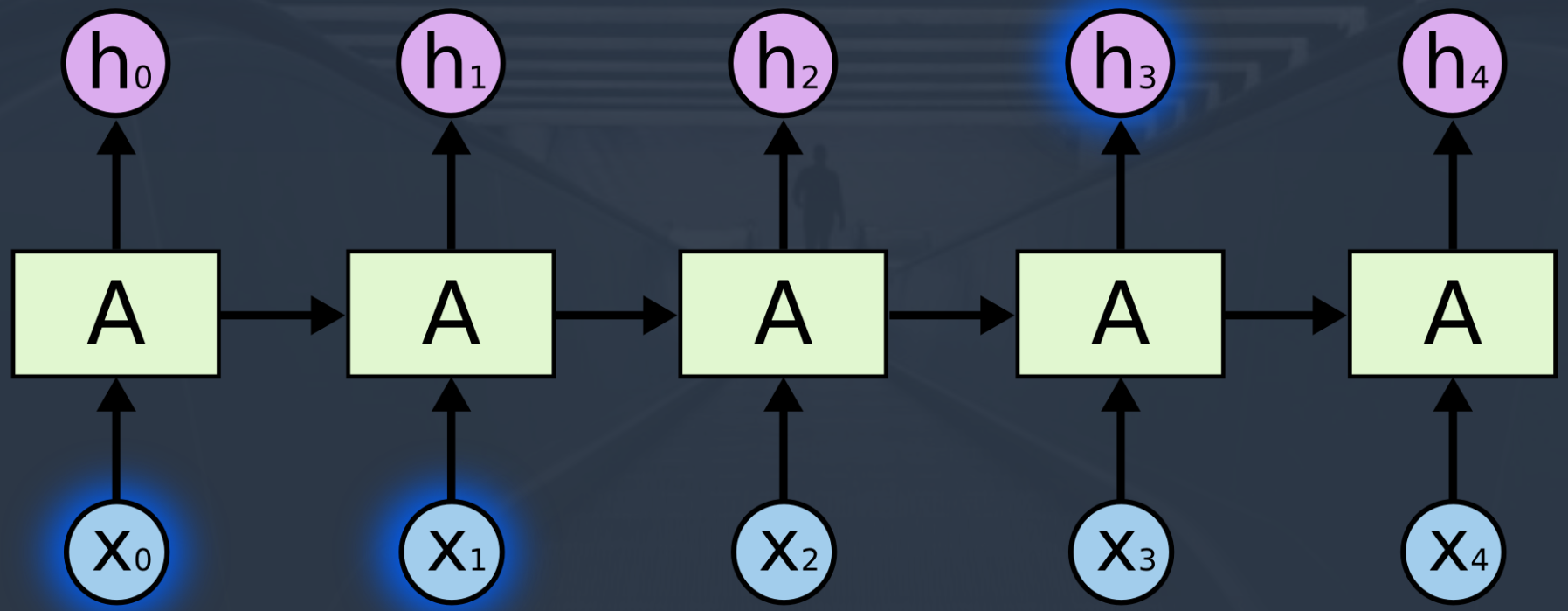
- دکتری هوش مصنوعی و رباتیکز
- مدیر عامل شرکت دانش بنیان ساعیان ارتباط

- هوش مصنوعی در مراکز فرودگاه‌های سراسر کشور، گروه صنعتی انتخاب
- مدیر پروژه سامانه‌های مخابراتی مترو تهران
- تولید نرم افزار شناسایی موضوع متون از با استفاده از روش های هوش مصنوعی
- تولید نرم افزار شناسایی علایم راهنمایی و رانندگی در خودروهای بدون سرنشین
- تولید Chatbot فارسی در حوزه سامانه‌های پروازی

- مدرس دوره‌های یادگیری ماشین و علم داده در همراه اول، وزارت نفت، بیمه ایران
- هم بنیان گذار Iran Machine Learning

RNN
(Recurrent Neural Network)





What is RNN?

Advantages of Recurrent Neural Network

1.RNN can model sequence of data so that each sample can be assumed to be dependent on previous ones

2.Recurrent neural network are even used with convolutional layers to extend the effective pixel neighborhood.

Disadvantages of Recurrent Neural Network

1.Gradient vanishing and exploding problems.

2.Training an RNN is a very difficult task.

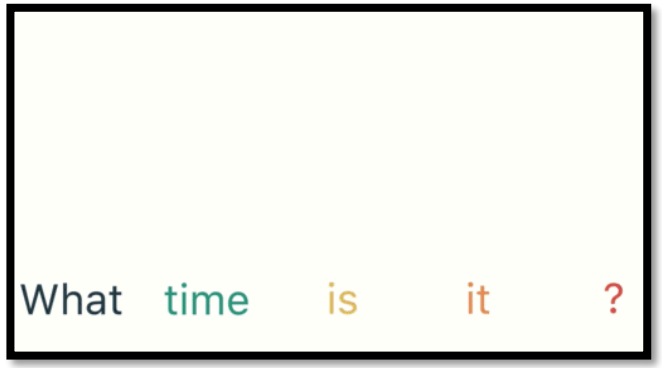
3.It cannot process very long sequences if using *tanh* or *relu* as an activation function.

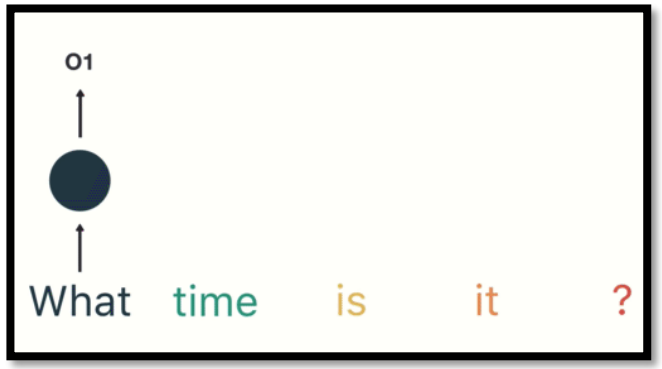
1. Unlike feedforward neural networks, RNNs can use their internal state (memory) to process sequences of inputs.

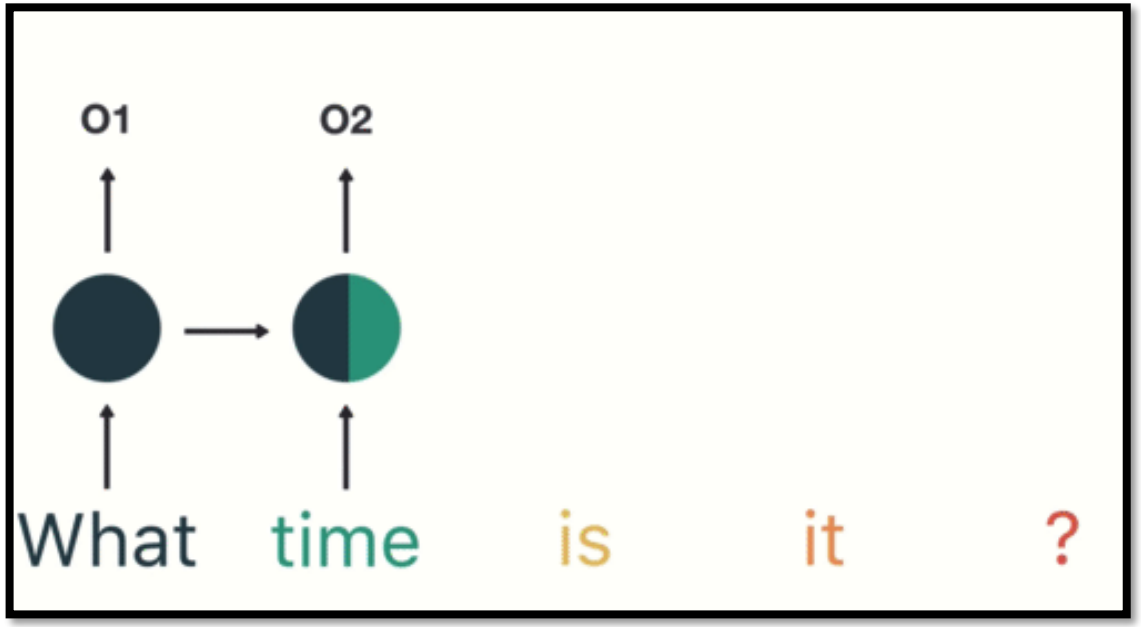
2. DNN is not for Sequential Data.

What is RNN?

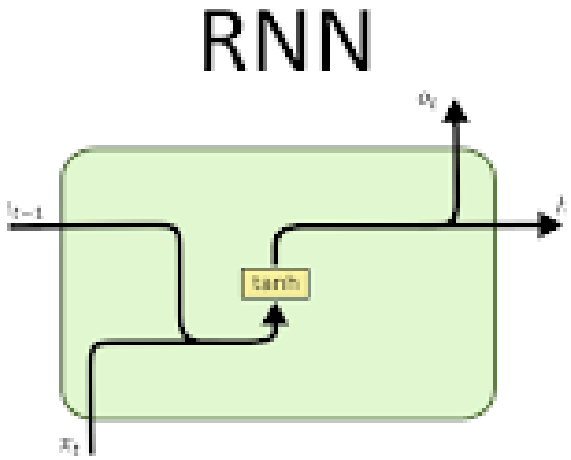
What time is it?

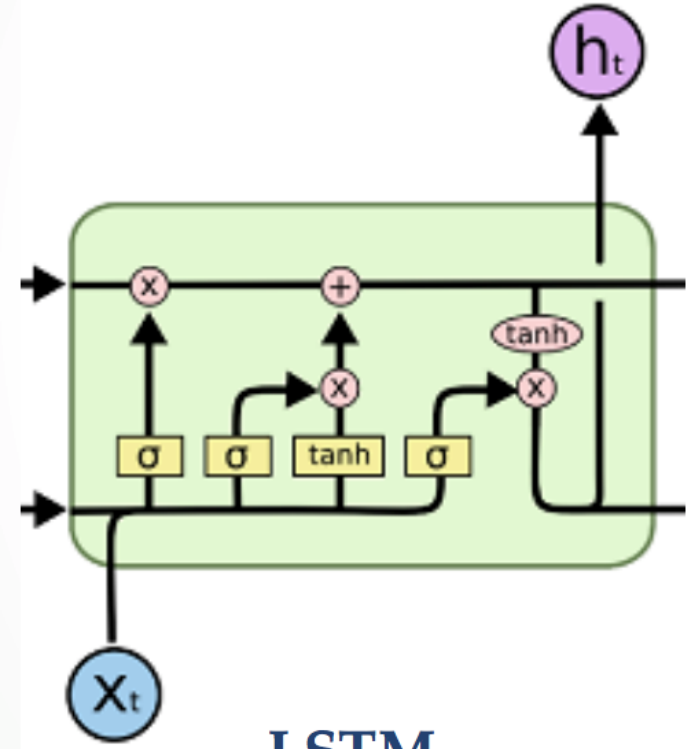

What
time
is
it
?


O1
What
time
is
it
?

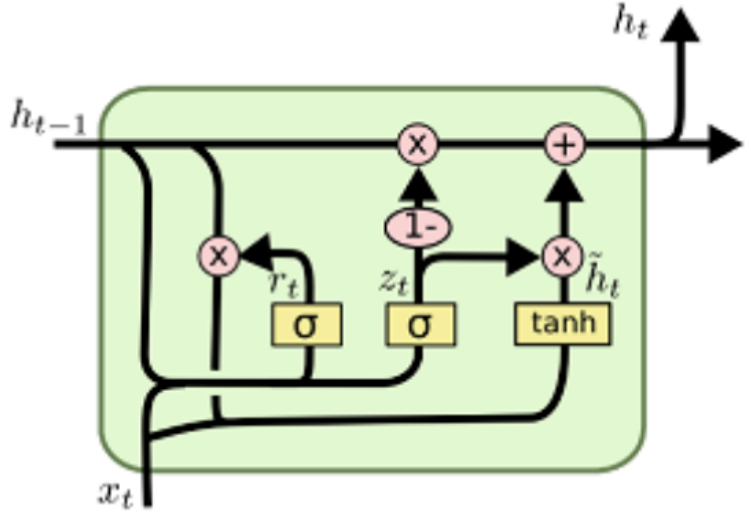

O1
O2
What
time
is
it
?

What is RNN?





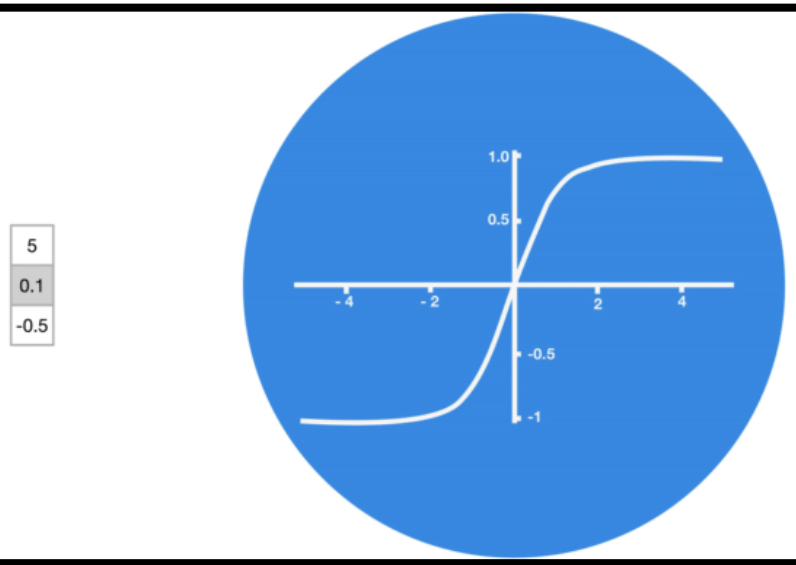
LSTM



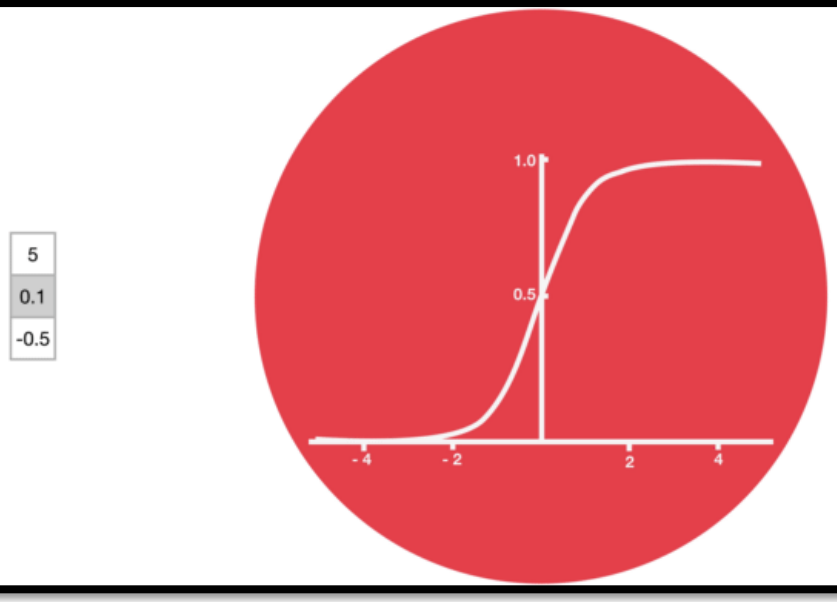
GRU

Tanh or Sigmoid?

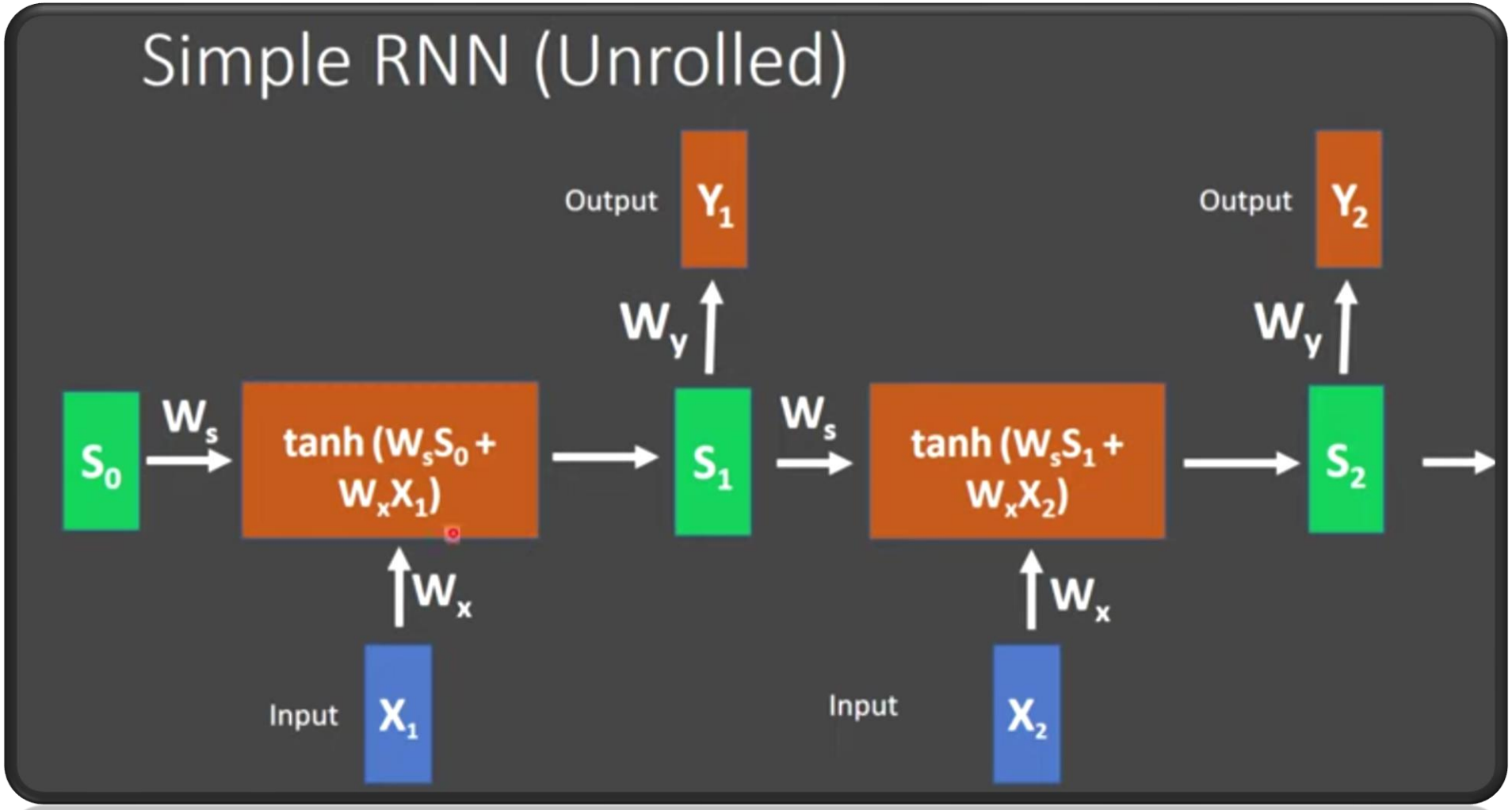
Tanh



Sigmoid



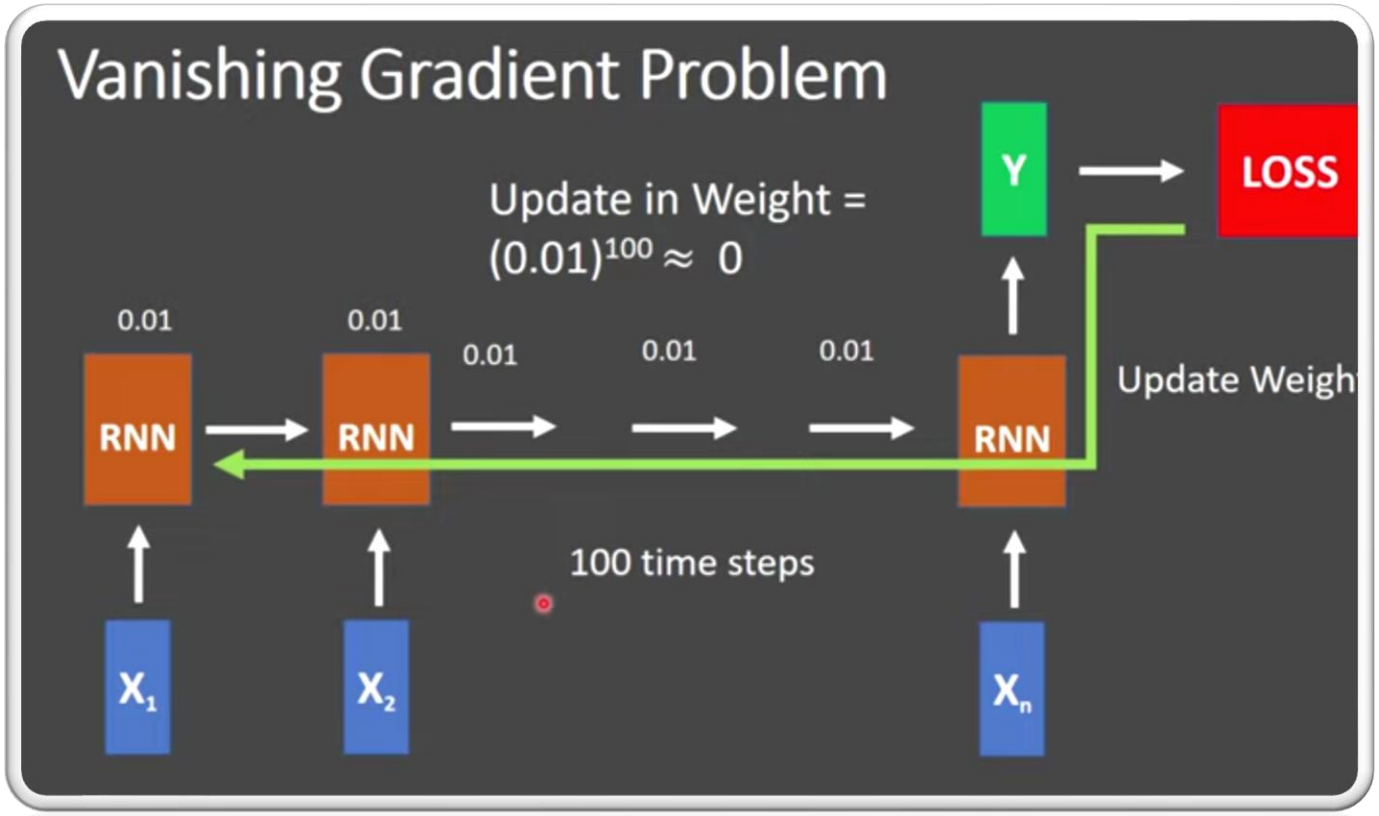
What is Simple RNN?

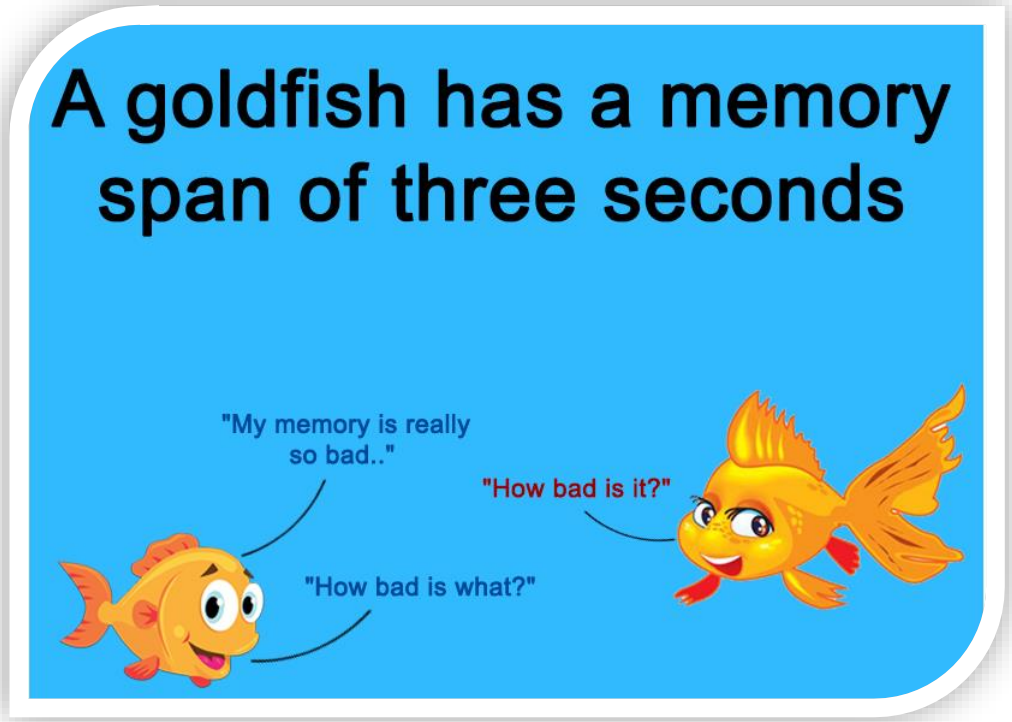


RNN Advantages ?

- RNN **Advantages**:
  - Can process any length input
  - Model size doesn't increase for longer input
  - Computation for step *t* can (in theory) use information from many steps back
  - Weights are shared across timesteps → representations are shared
- RNN **Disadvantages**:
  - Recurrent computation is slow
  - In practice, difficult to access information from many steps back

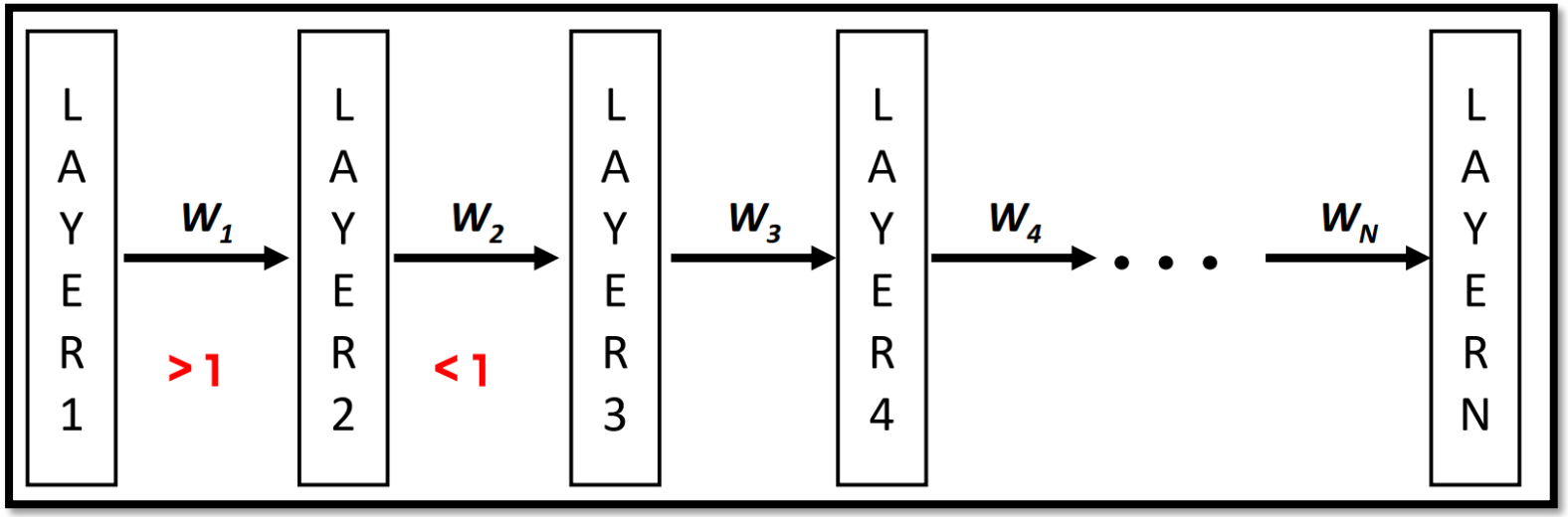
Simple RNN Disadvantages?



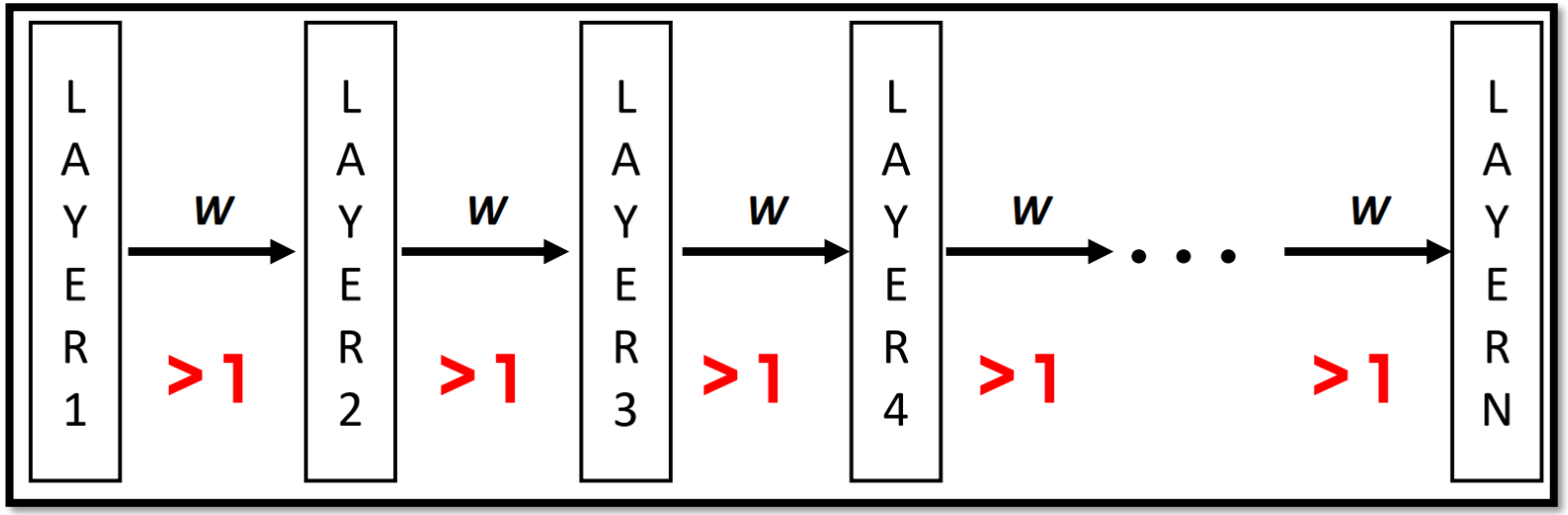


How to Deal with Vanishing/Exploding Gradient?

DNN (Deep Neural Network)



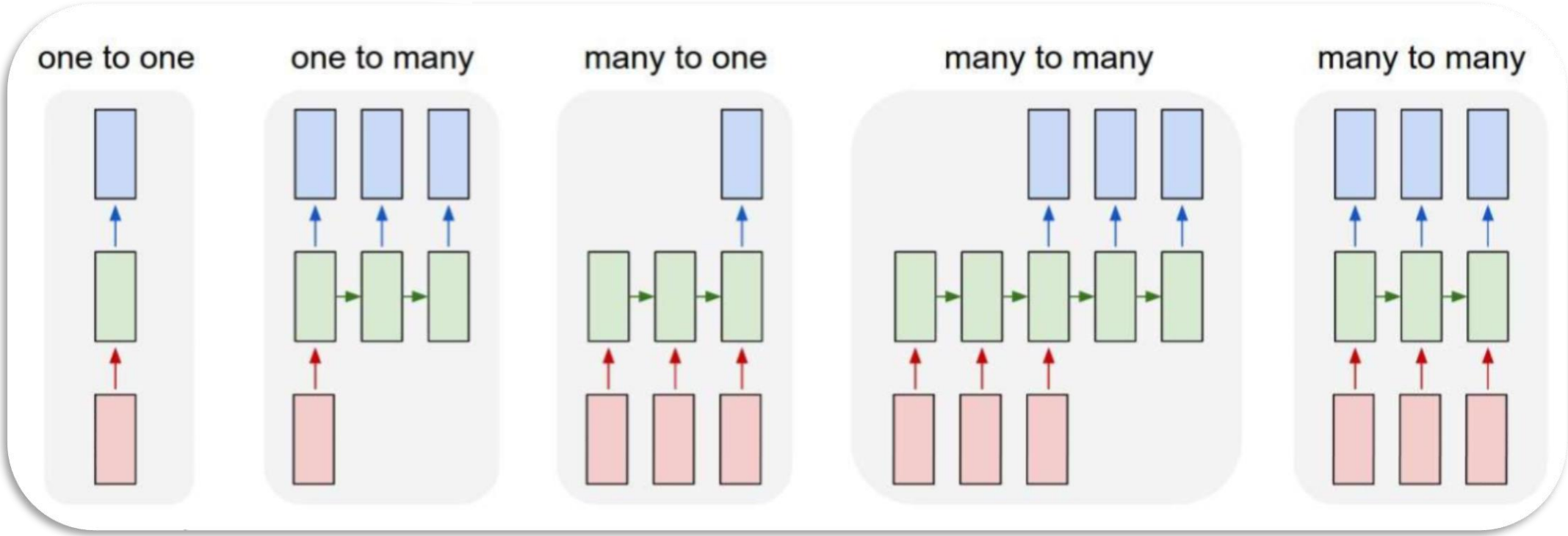
RNN (Recurrent Neural Network)



Different kind of recurrent networks?

Image Captioning
image -> sequence of words

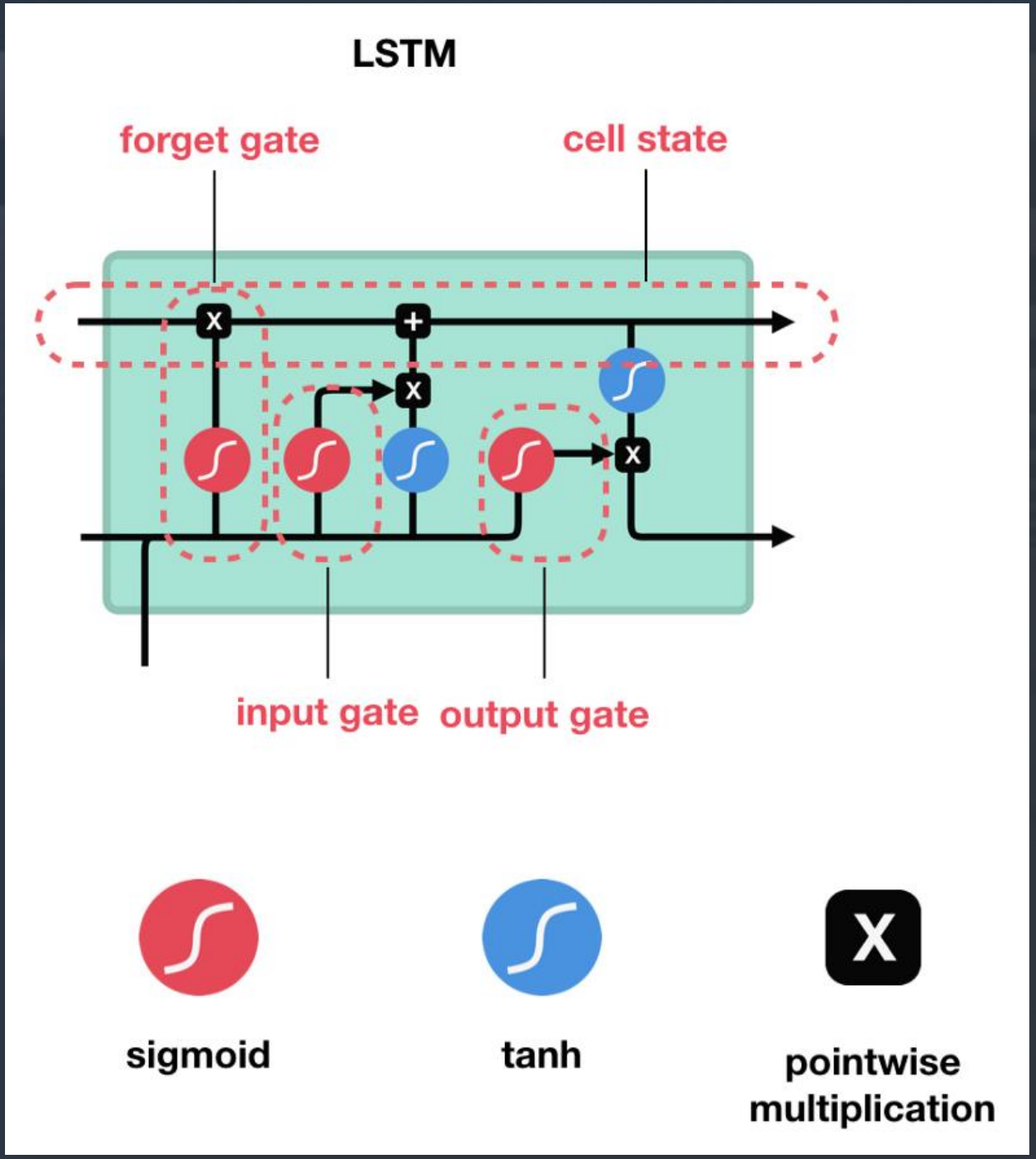
Machine Translation
seq of words -> seq of words



Vanilla Neural Networks

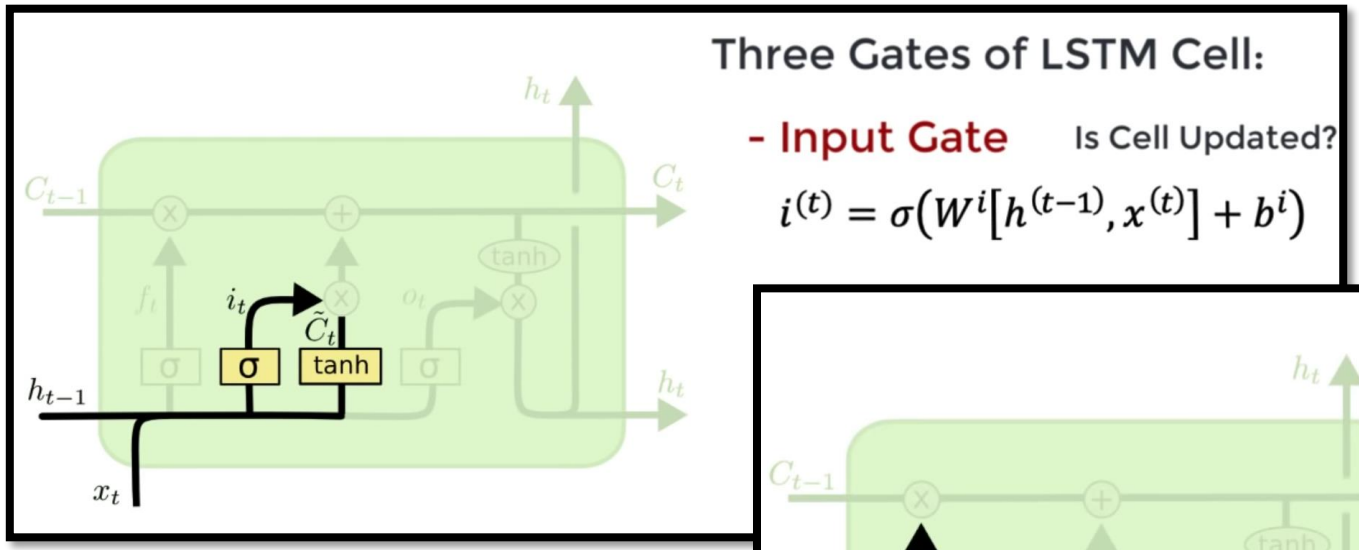
Sentiment Classification
sequence of words -> sentiment

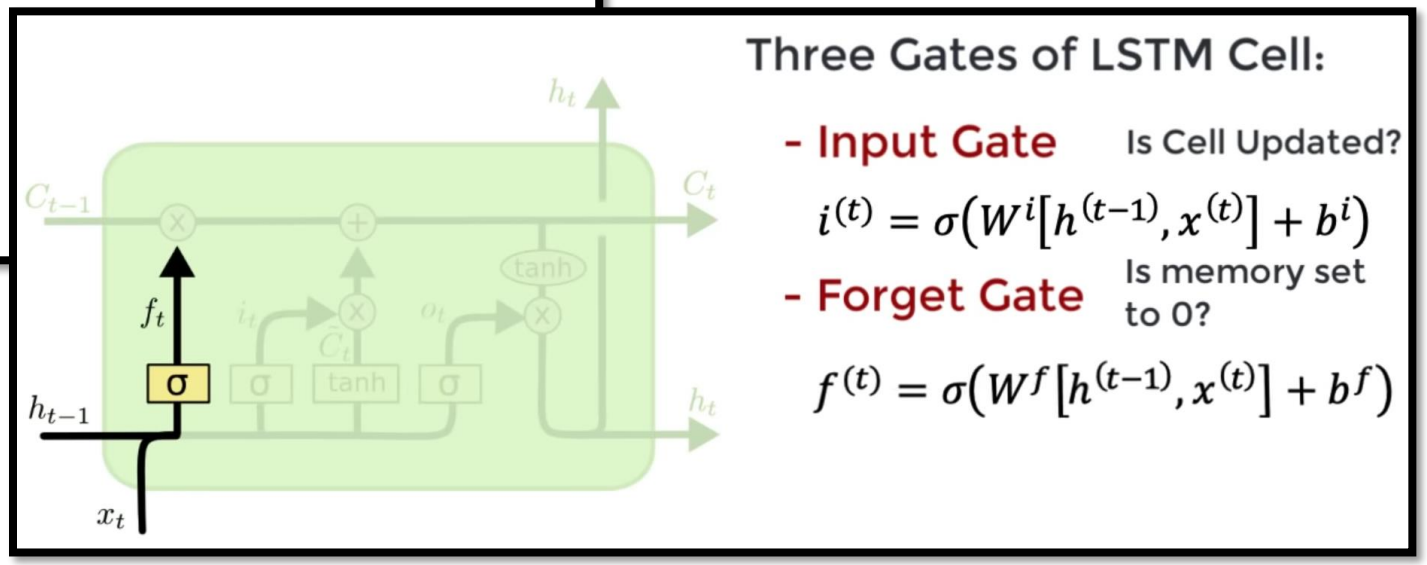
Video classification on frame level

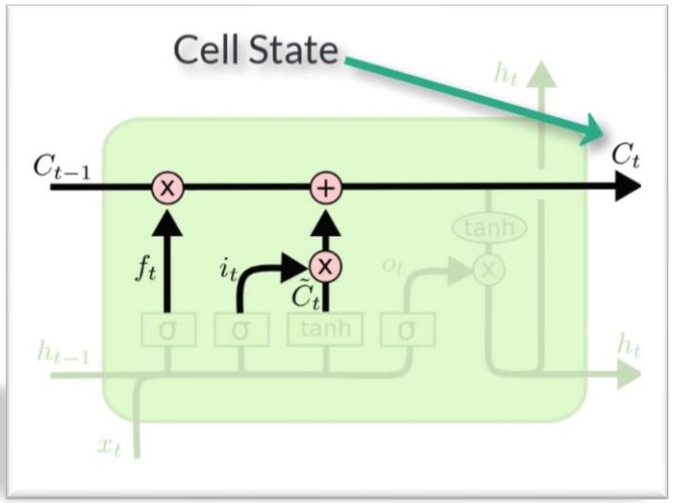


LSTM
(Long Short Term Memory)

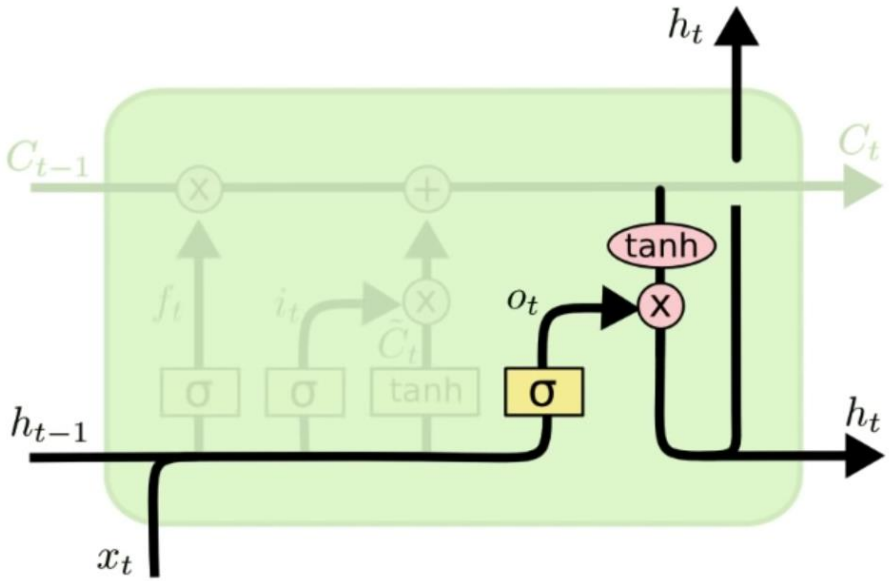
What is LSTM?


Three Gates of LSTM Cell:
- Input Gate
Is Cell Updated?
$i^{(t)} = \sigma(W^i[h^{(t-1)}, x^{(t)}] + b^i)$
h_t
C_{t-1}
C_t
f_t
i_t
$\tilde{C}_t$
o_t
h_{t-1}
x_t
σ
tanh


Three Gates of LSTM Cell:
- Input Gate
Is Cell Updated?
$i^{(t)} = \sigma(W^i[h^{(t-1)}, x^{(t)}] + b^i)$
- Forget Gate
Is memory set to 0?
$f^{(t)} = \sigma(W^f[h^{(t-1)}, x^{(t)}] + b^f)$
h_t
C_{t-1}
C_t
f_t
i_t
$\tilde{C}_t$
o_t
h_{t-1}
x_t
σ
tanh


Cell State
h_t
C_{t-1}
C_t
f_t
i_t
$\tilde{C}_t$
o_t
h_{t-1}
x_t
σ
tanh

What is LSTM?



Three Gates of LSTM Cell:

- **Input Gate** — Is Cell Updated?

$$i^{(t)} = \sigma\left(W^i\left[h^{(t-1)}, x^{(t)}\right] + b^i\right)$$

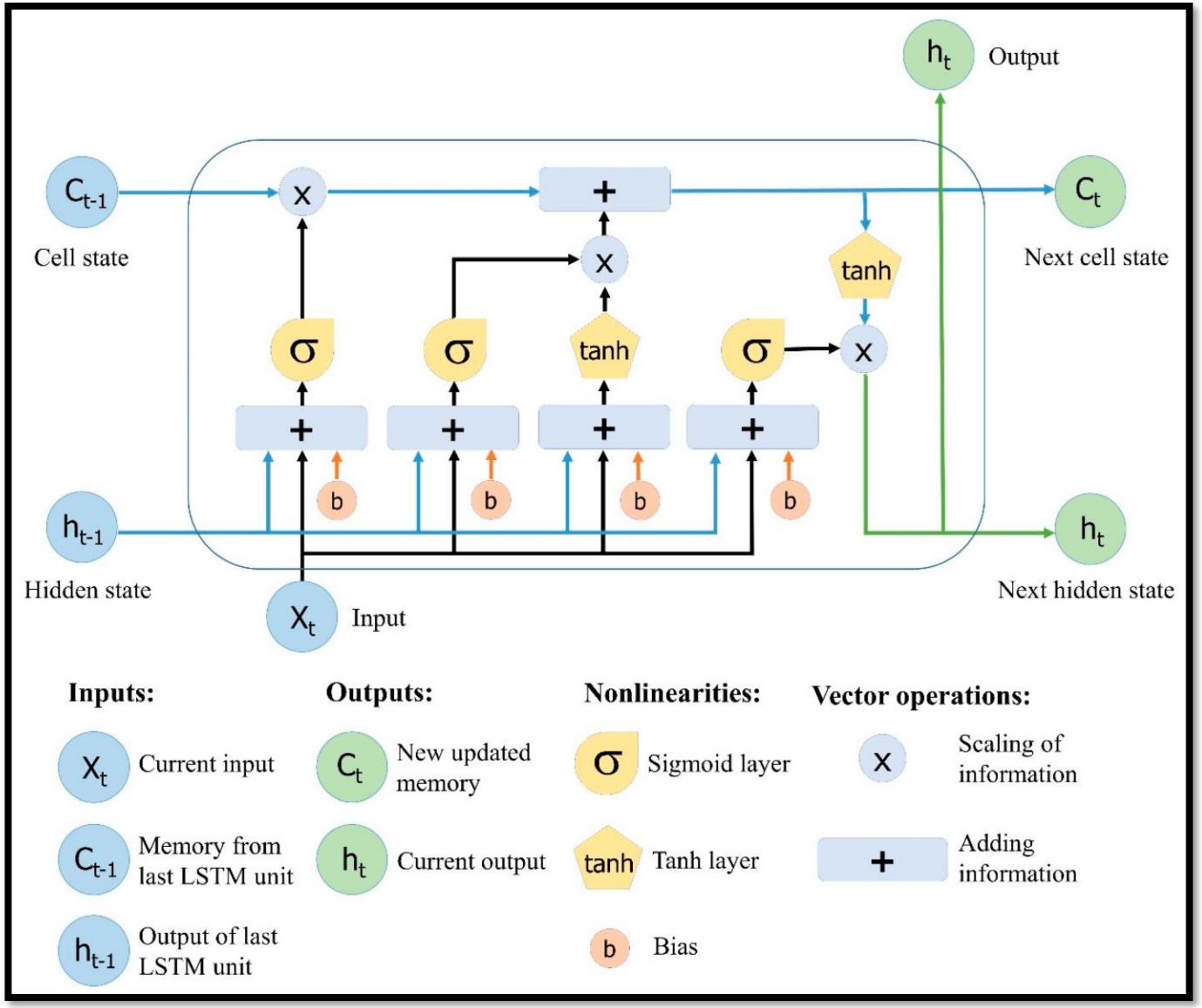
- **Forget Gate** — Is memory set to 0?

$$f^{(t)} = \sigma\left(W^f\left[h^{(t-1)}, x^{(t)}\right] + b^f\right)$$

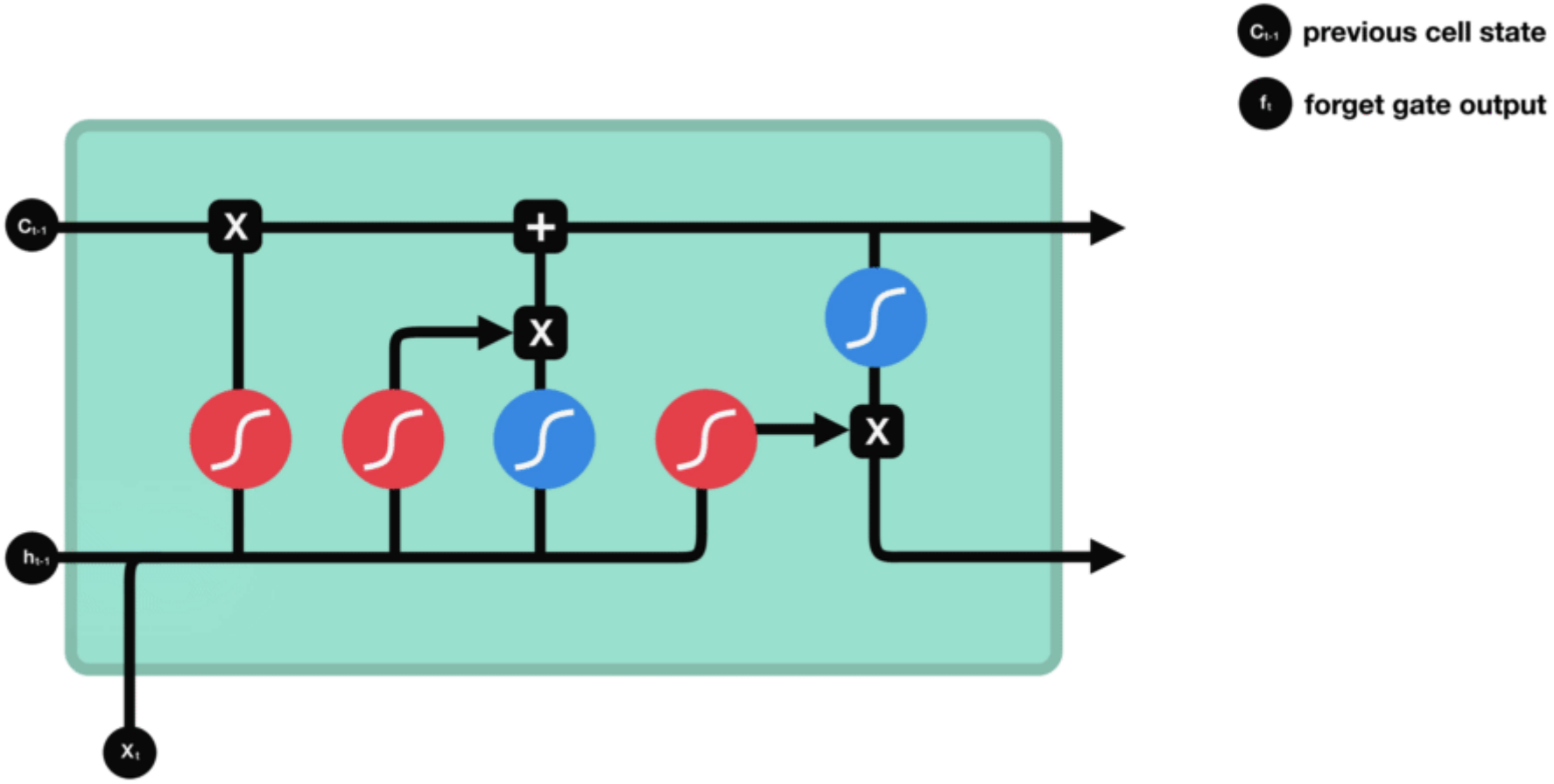
- **Output Gate** — Is current info visible?

$$o^{(t)} = \sigma\left(W^o\left[h^{(t-1)}, x^{(t)}\right] + b^o\right)$$

What is LSTM?

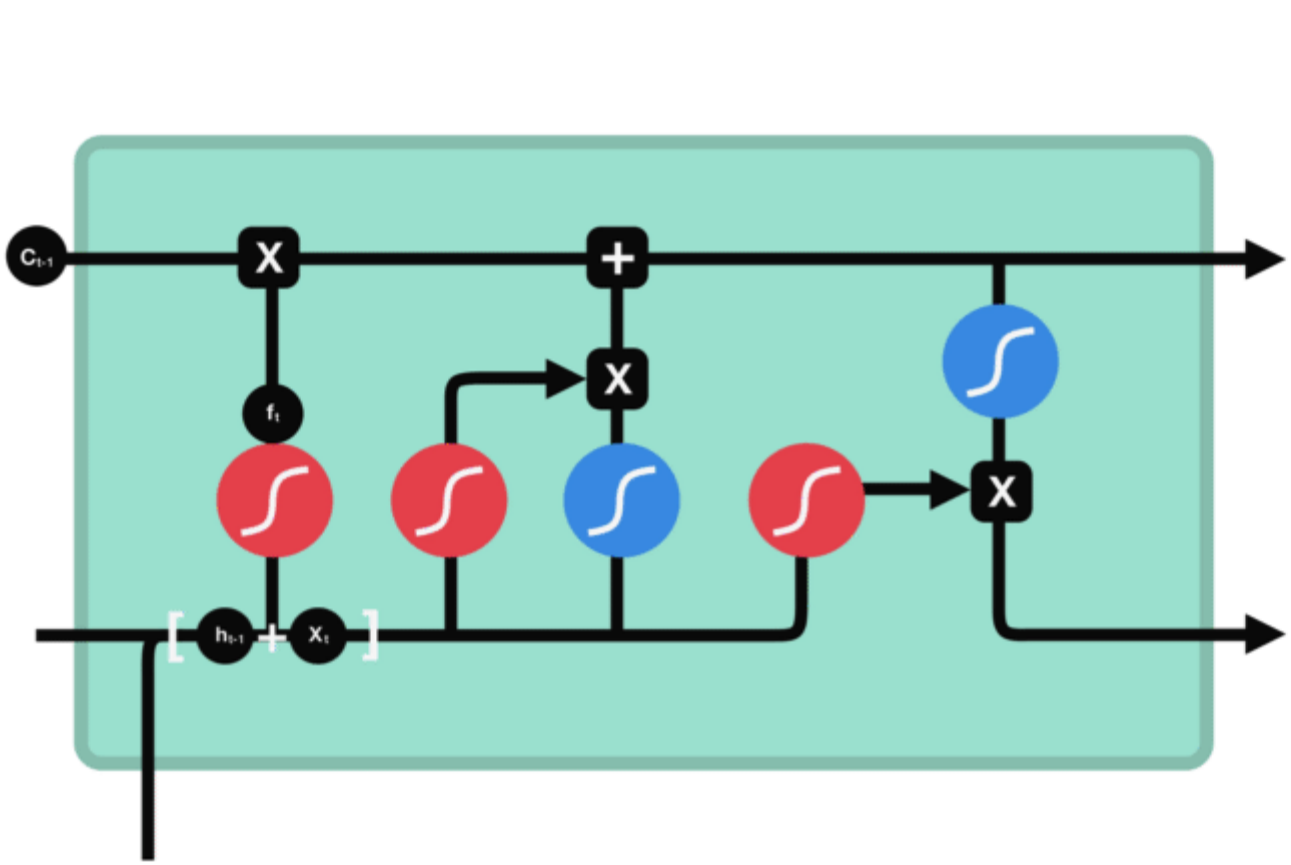

h_t Output
C_{t-1} Cell state
C_t Next cell state
tanh
σ
tanh
σ
h_{t-1} Hidden state
h_t Next hidden state
b
X_t Input
Inputs:
X_t Current input
C_{t-1} Memory from last LSTM unit
h_{t-1} Output of last LSTM unit
Outputs:
C_t New updated memory
h_t Current output
Nonlinearities:
σ Sigmoid layer
tanh Tanh layer
b Bias
Vector operations:
Scaling of information
Adding information

What is LSTM? (1)



Forget gate operations

What is LSTM? (2)



C_{t-1} previous cell state

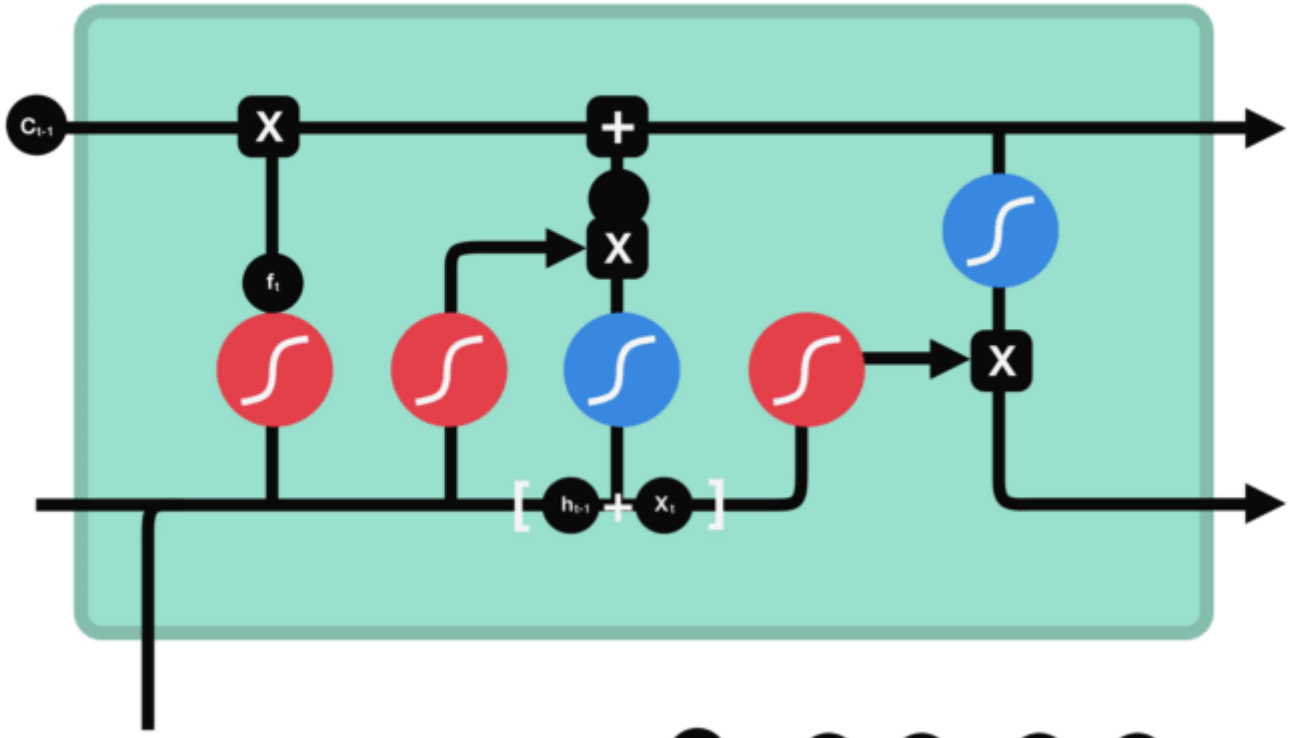
f_t forget gate output

i_t input gate output

$\tilde{C}_t$ candidate

Input gate operations

What is LSTM? (3)



C_{t-1} previous cell state

f_t forget gate output

i_t input gate output

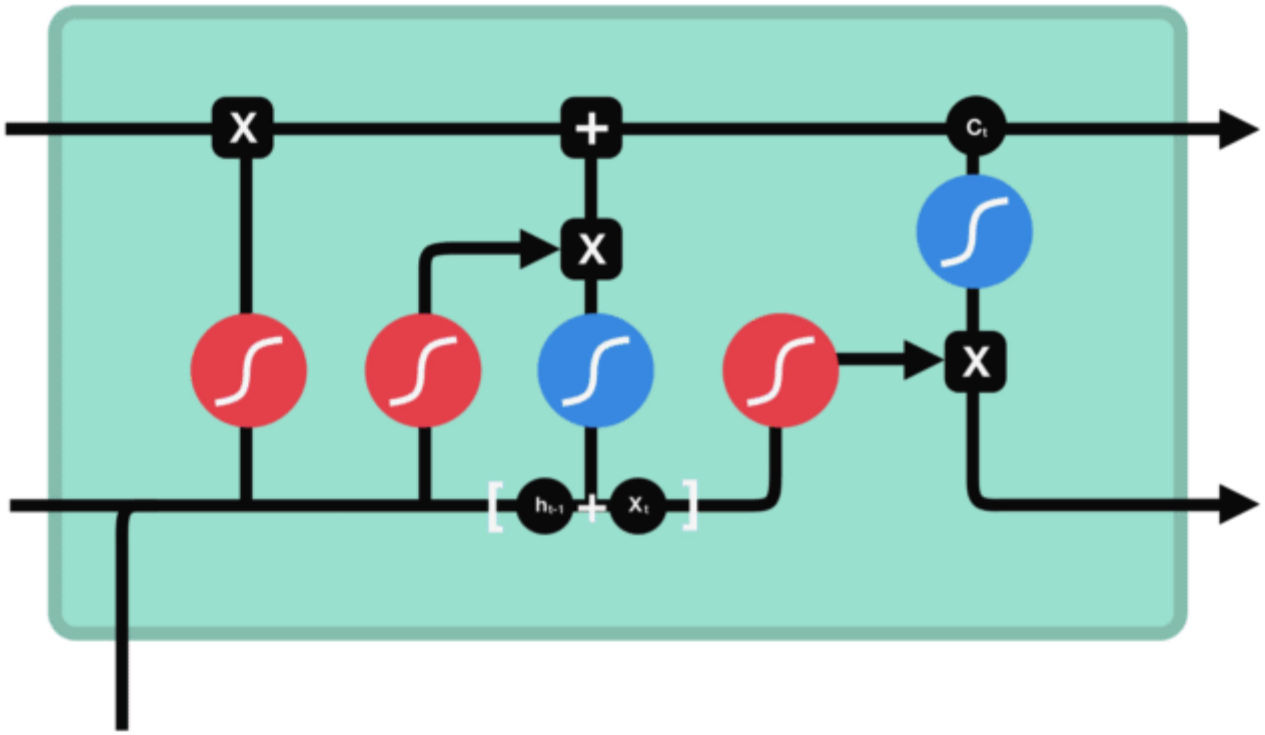
$\tilde{C}_t$ candidate

C_t new cell state

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

Calculating cell state

What is LSTM? (4)



c_{t-1} previous cell state

f_t forget gate output

i_t input gate output

$\tilde{C}_t$ candidate

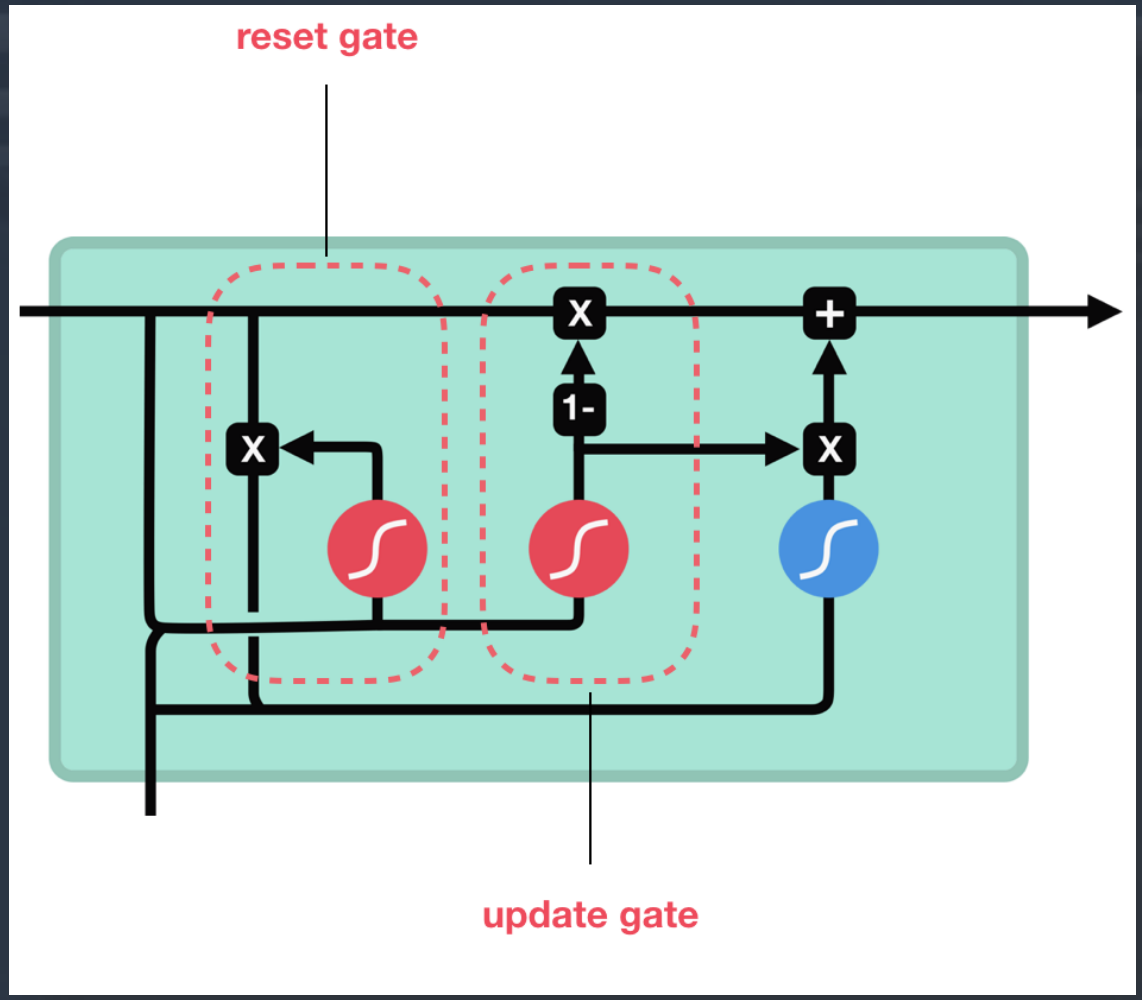
C_t new cell state

o_t output gate output

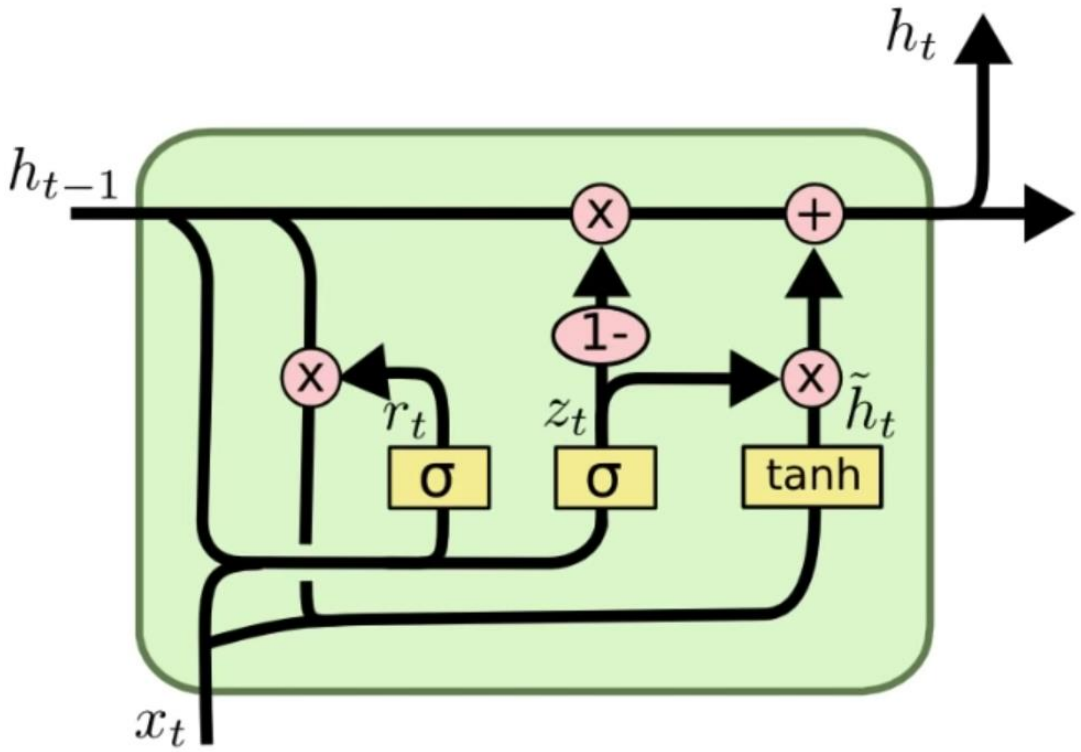
h_t hidden state

output gate operations

GRU
(Gated Recurrent Unit)



What is LSTM?



TWO Gates of GRU Cell:

- Update Gate

$$z^{(t)} = \sigma\left(W^z\left[h^{(t)}, x^{(t)}\right] + b^z\right)$$

- Reset Gate

$$r^{(t)} = \sigma\left(W^r\left[h^{(t)}, x^{(t)}\right] + b^r\right)$$

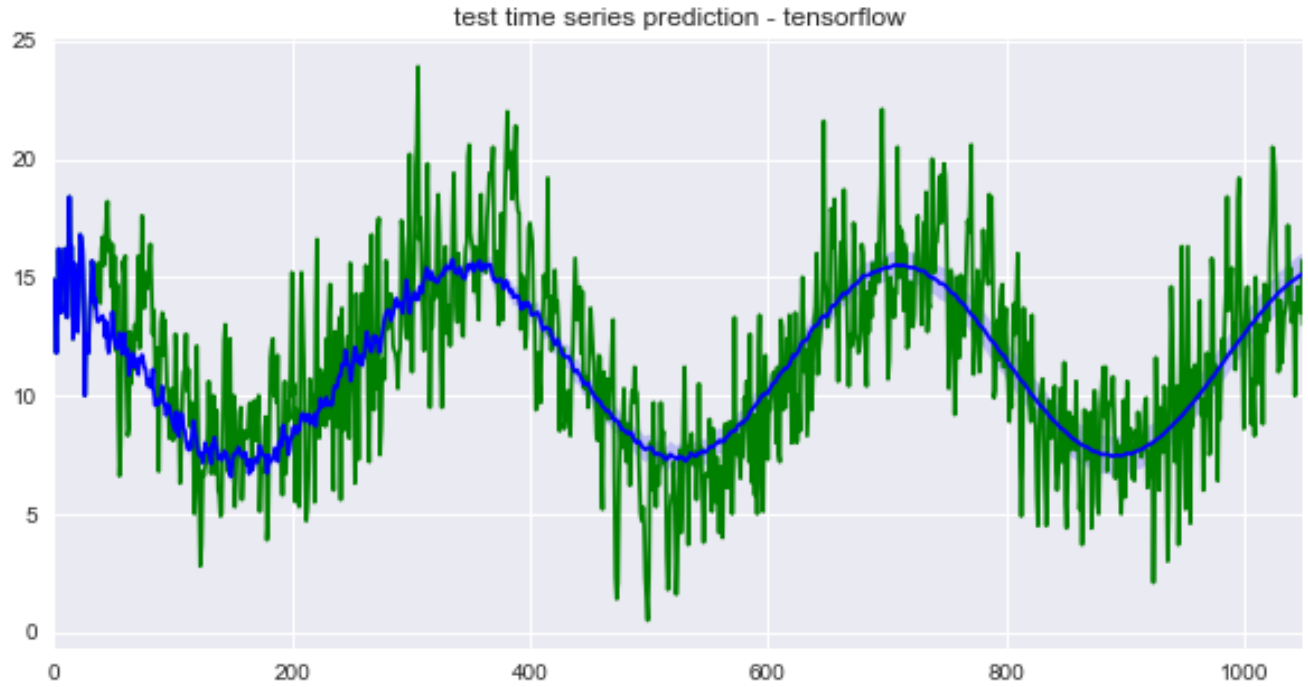
Time Series Analysis with LSTM

What is Autoregression?

A time series is a sequence of measurements of the same variable(s) made over time.

An **autoregressive model** is when a value from a time series is regressed on previous values from that same time series. for example, y_t on y_{t-1}:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \epsilon_t.$$



Different kind of Time-Series dataset (Multivariate Time Series)

The number of observations recorded for a given time in a time series dataset matters. Traditionally, different names are used:

- **Univariate Time Series**: These are datasets where only a single variable is observed at each time, such as temperature each hour.

- **Multivariate Time Series**: These are datasets where two or more variables are observed at each time.

Most time series analysis methods, and even books on the topic, focus on univariate data. This is because it is the simplest to understand and work with.
Multivariate data is often more difficult to work with. It is harder to model and often many of the classical methods do not perform well.

Different kind of Time-Series dataset (Multi-Step Forecasting)

The number of time steps ahead to be forecasted is important.
Again, it is traditional to use different names for the problem depending on the number of time-steps to forecast:

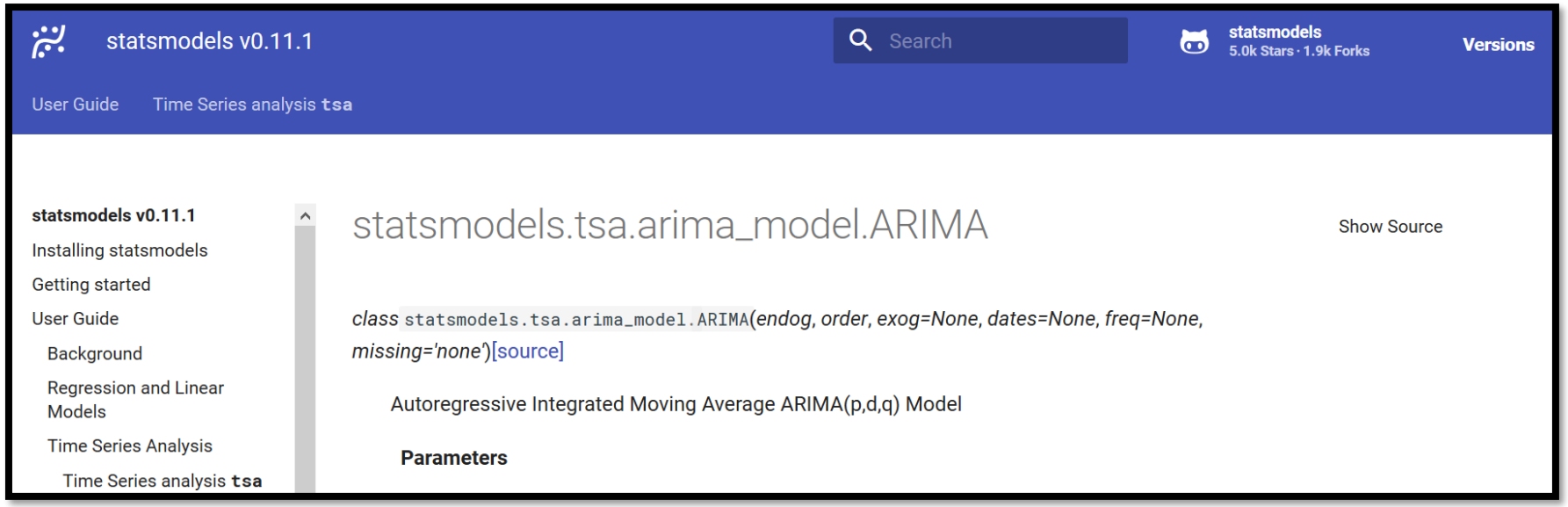
- **One-Step Forecast**: This is where the next time step (t+1) is predicted.
- **Multi-Step Forecast**: This is where two or more future time steps are to be predicted.

Time Series Forecasting as Supervised Learning

Solutions to Time-Series

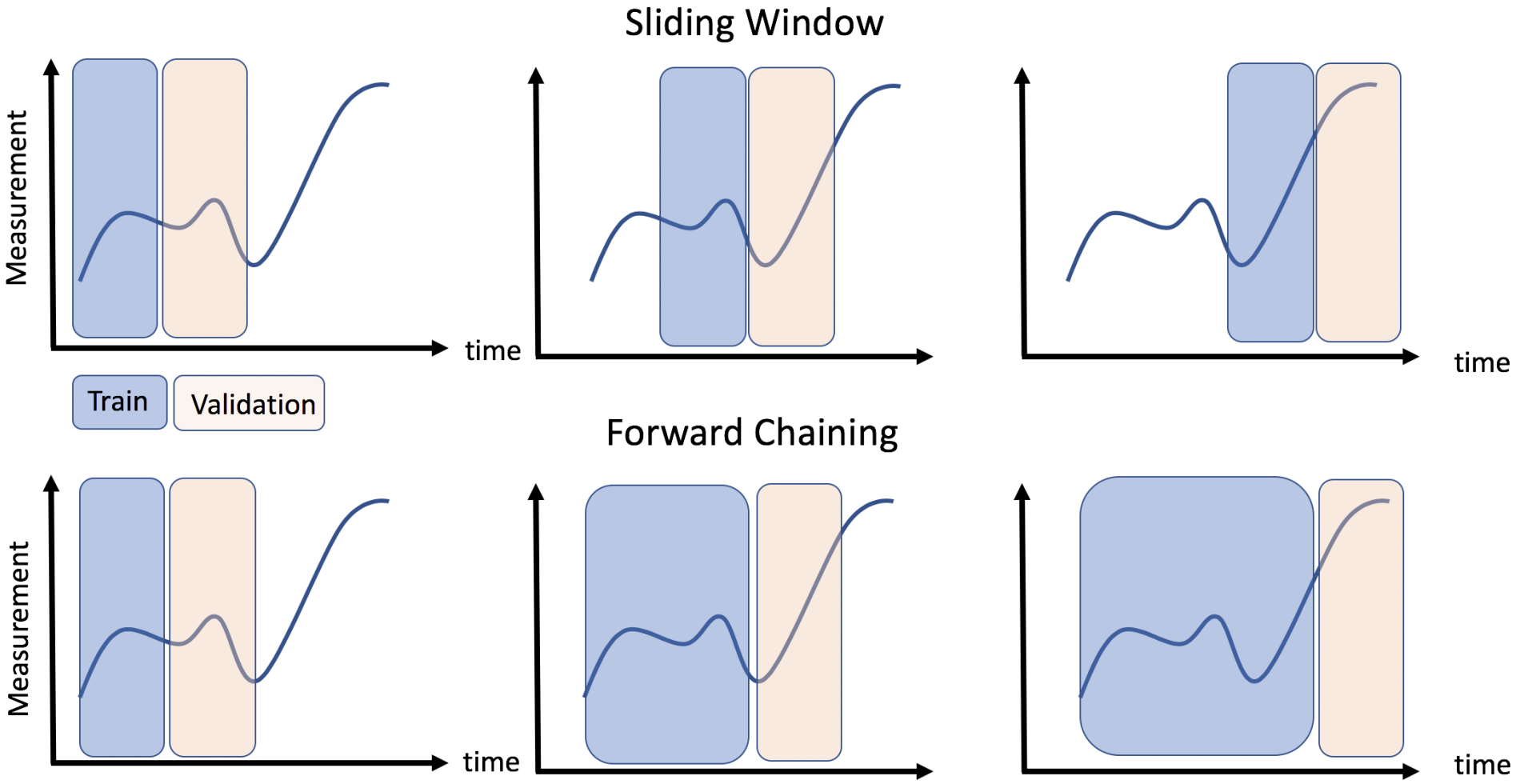
What is Autoregression solutions?

ARIMA, short for **'Auto Regressive Integrated Moving Average'** is actually a class of models that 'explains' a given time series based on its own past values, that is, its own lags and the lagged forecast errors, so that equation can be used to forecast future values.



What is Autoregression solutions?

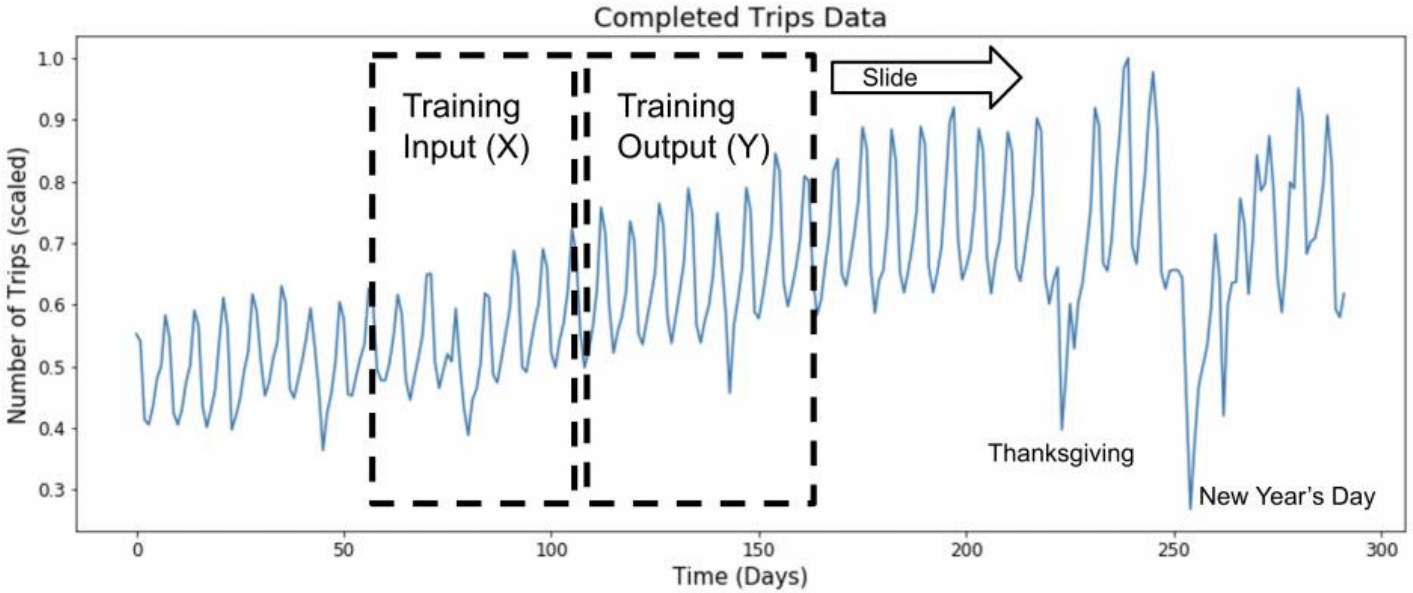
Sliding window vs. forward chaining cross validation



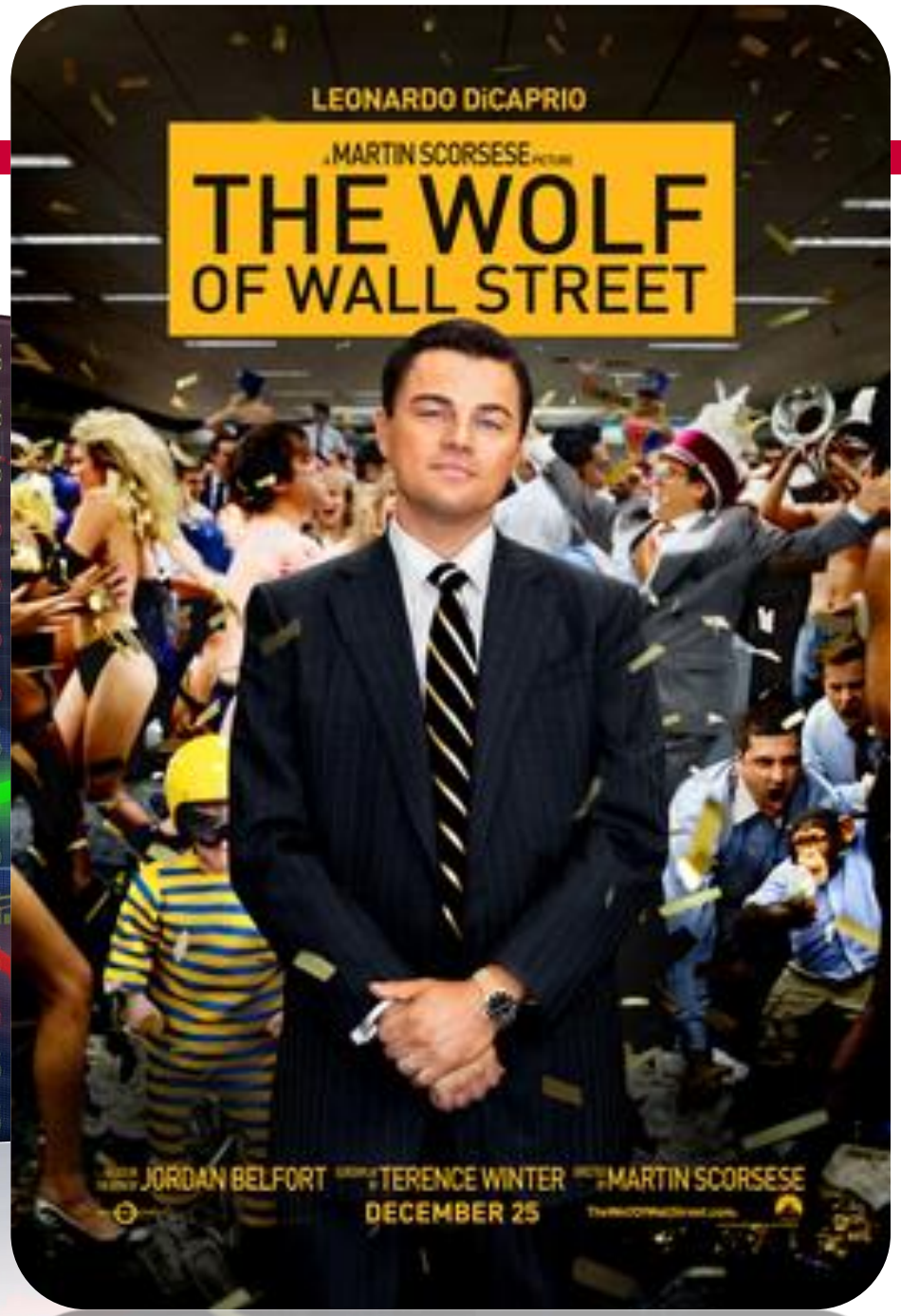
LSTM Model Architecture for Rare Event Time Series Forecasting

A training dataset was created by splitting the historical data into sliding windows of input and output variables.

The specific size of the look-back and forecast horizon used in the experiments were not specified in the paper.



Scaled Multivariate Input for Model Taken from “Time-series Extreme Event Forecasting with Neural Networks at Uber”.



Just Click Here

Get Your Hands Dirty

ممنون
Dankie
Gracias
Спасибо
شكراً
Merci
Takk
Köszönjük
Terima kasih
Grazie
Dziękujemy
Dėkojame
Ďakujeme
Vielen Dank
Paldies
Kiitos
Täname teid
谢谢
Thank You
Tak
感謝您
Obrigado
Teşekkür Ederiz
감사합니다
Σας Ευχαριστούμ
ขอบคุณ
Bedankt
Děkujeme vám
ありがとうございます
Tack