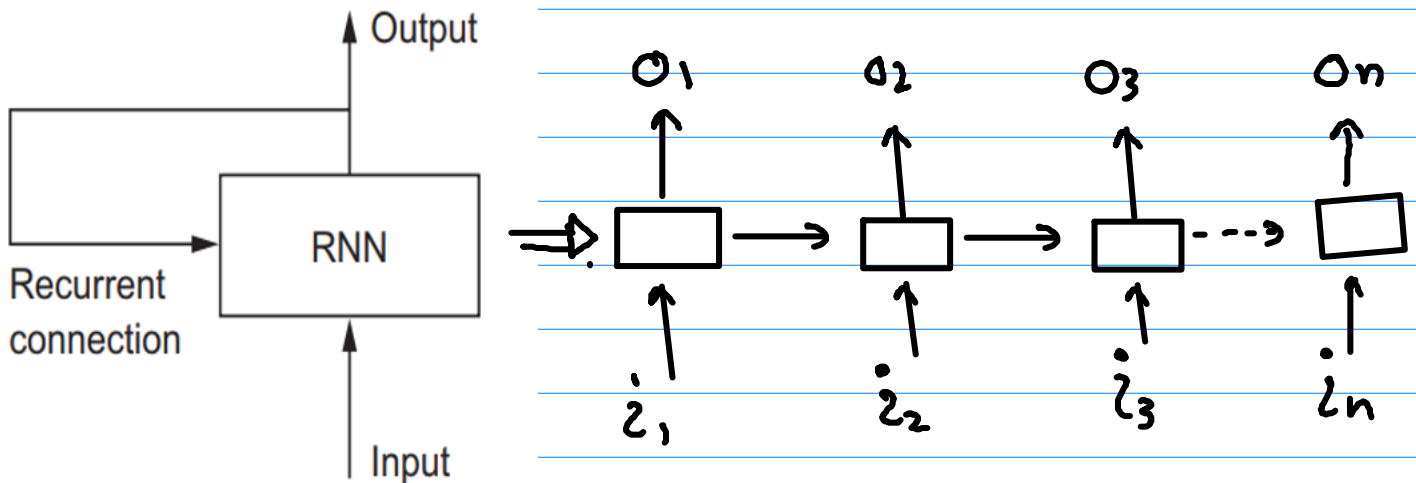
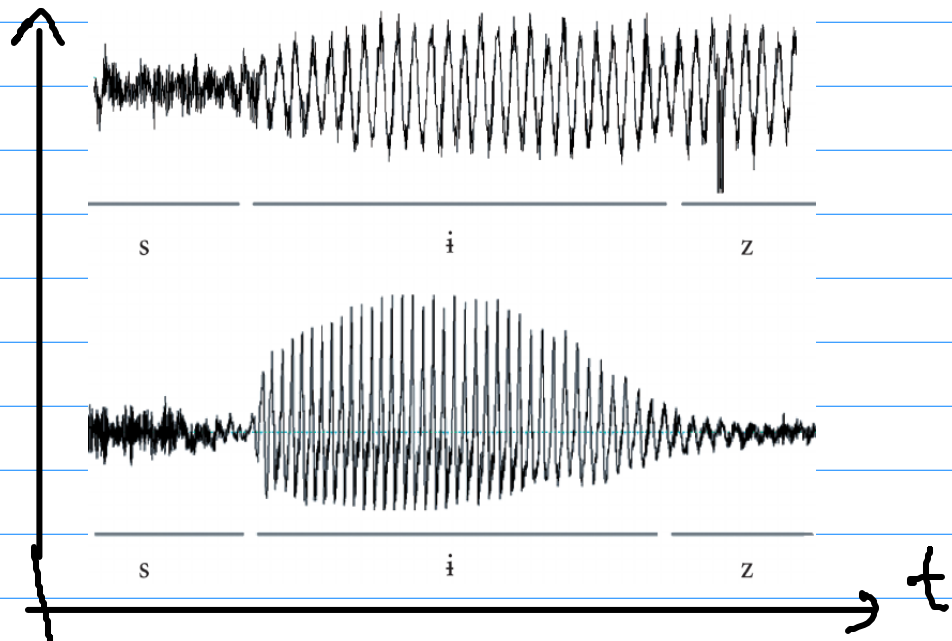


## Recurrent Neural Networks

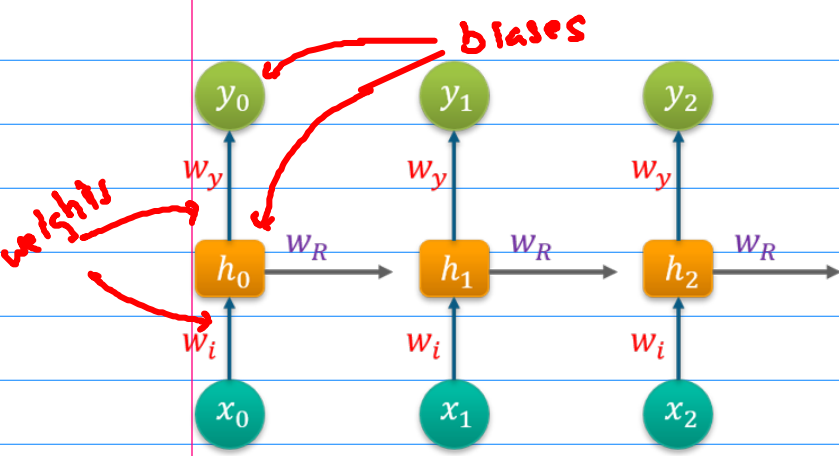


## Sound wave

A



# Math Behind RNN



$x$  - inputs  $y$  - outputs

$h$  - intermediate value

$w$  - weights,  $b$  - biases

Function (ex. tanh)

$$h^{(t)} = g_h(w_i x^{(t)} + w_R h^{(t-1)} + b_h)$$

$$y^{(t)} = g_y(w_y h^{(t)} + b_y)$$

Function

$t=0$ .  $h_0 = g_h(w_i x_0 + w_R \cdot 0 + b_h) = g_h(w_i x_0 + b_h)$   
 $y_0 = g_y(w_y h_0 + b_y)$   $y_0$  - depend on  $x_0$

$t=1$ .  $h_1 = g_h(w_i x_1 + w_R \cdot h_0 + b_h)$   
 $y_1 = g_y(w_y h_1 + b_y)$   $y_1$  - depend on  $x_0, x_1$

$t=2$ .  $h_2 = g_h(w_i x_2 + w_R h_1 + b_h)$   
 $y_2 = g_y(w_y h_2 + b_y)$   $y_2$  - depend on  $x_0, x_1, x_2$

Sigmoid function

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f)$$

$w_f$  = Weight

$h_{t-1}$  = Output from previous timestamp

$x_t$  = New input

$b_f$  = Bias

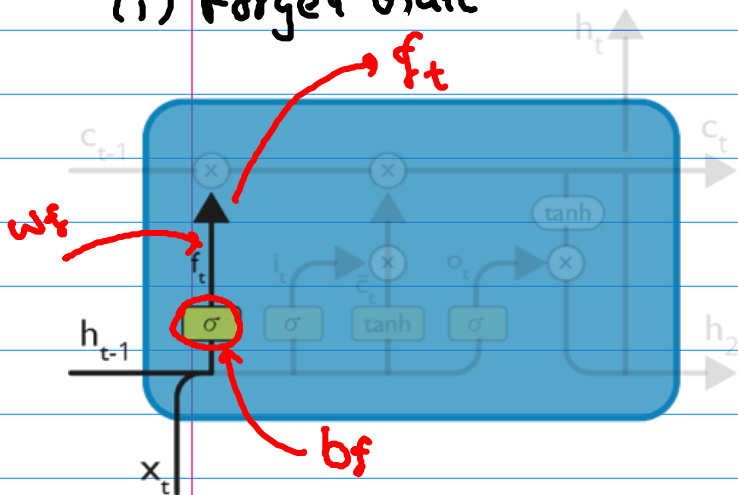
$$0 \leq f_t \leq 1$$

\*  $f_t = 1$  (Keep the previous state in memory)

\*  $f_t = 0$  (trash it)

## LSTM

(1) Forget Gate.



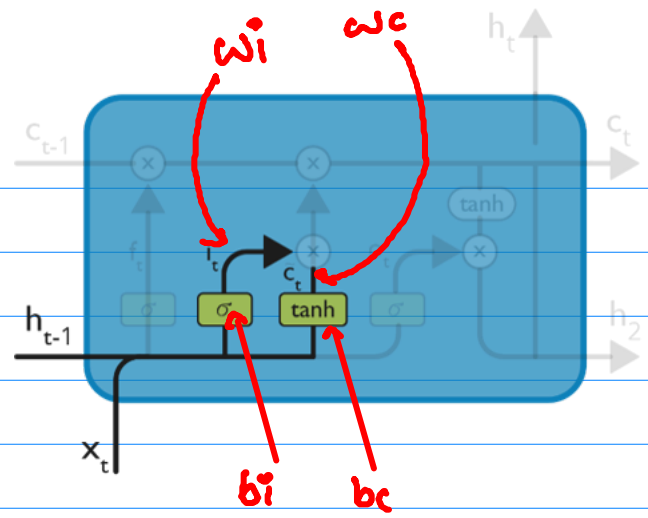
## (2) Input Gate Layer

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i)$$

*σ sigmoid*

$$\tilde{c}_t = \tanh(w_c[h_{t-1}, x_t] + b_c)$$

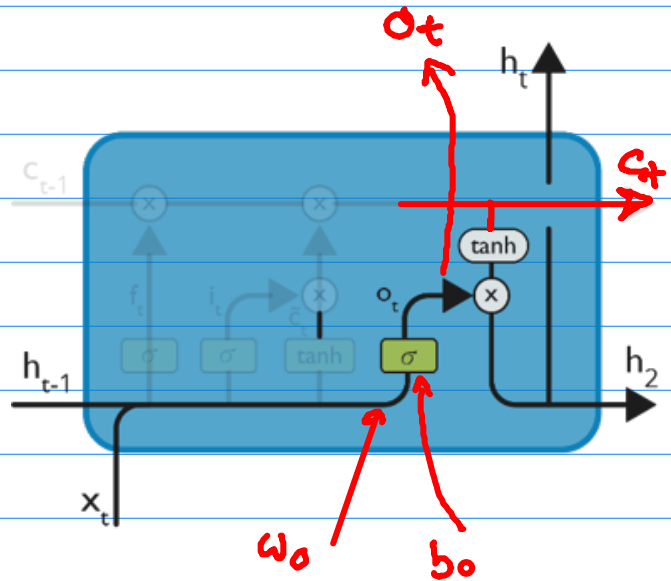
*hyperbolic tan*



## (3) Output Gate

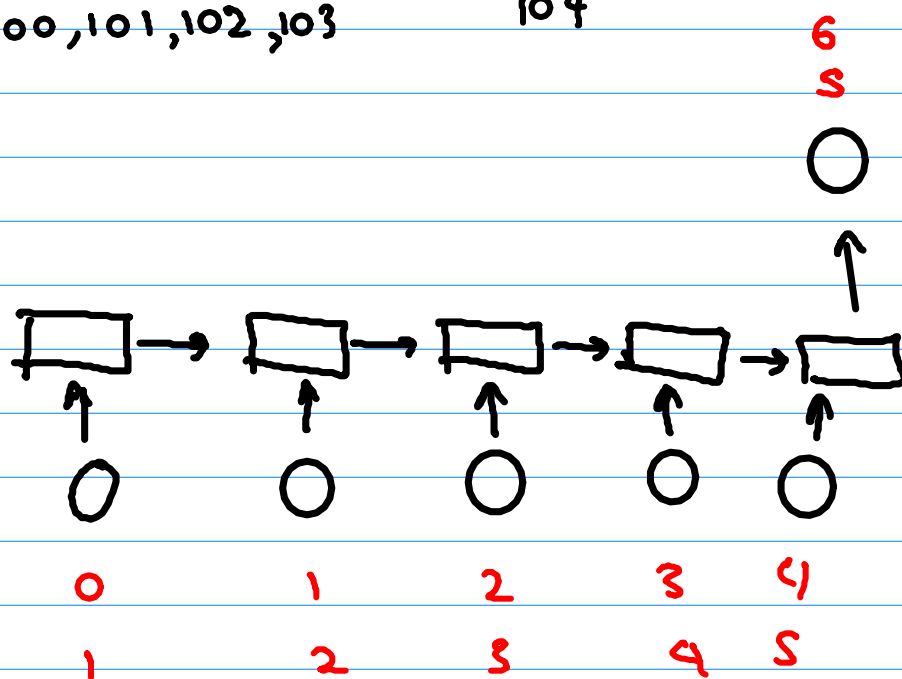
$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(c_t)$$



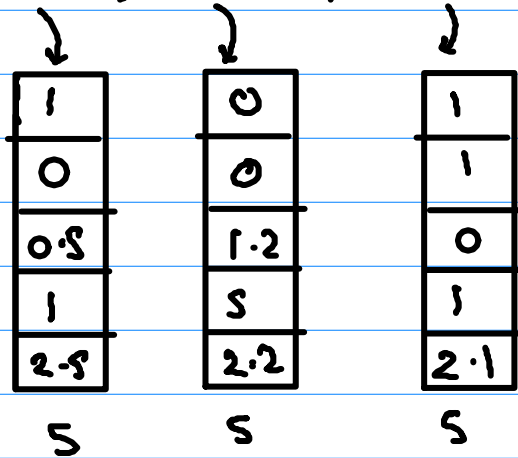
# 1.0 RNN For a Simple problem

data	target
0, 1, 2, 3, 4	5
1, 2, 3, 4, 5	6
2, 3, 4, 5, 6	7
...	
99, 100, 101, 102, 103	104

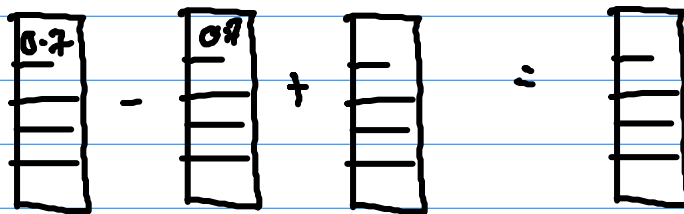


## Word 2 Vec

mother, father, Brother  $\leftarrow$  vocabulary



King - Man + Women = Queen



France - Paris + Rome = Italy

## Chat-Bot Plan

