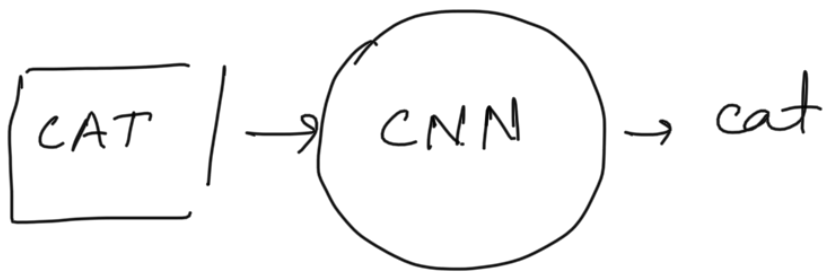
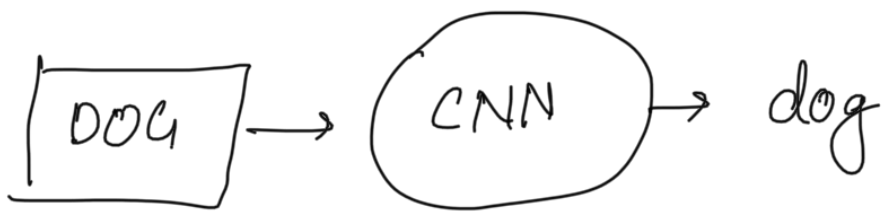
RNN

Q. Why not Feed forward networks / CNN ?

"In FF networks, current inputs are not related on the previous inputs"

For eg





The ans to this input does not depend on the previous input(cat) or previous outputs of the netwok

↓ BUT

sometimes we have data that is sequential

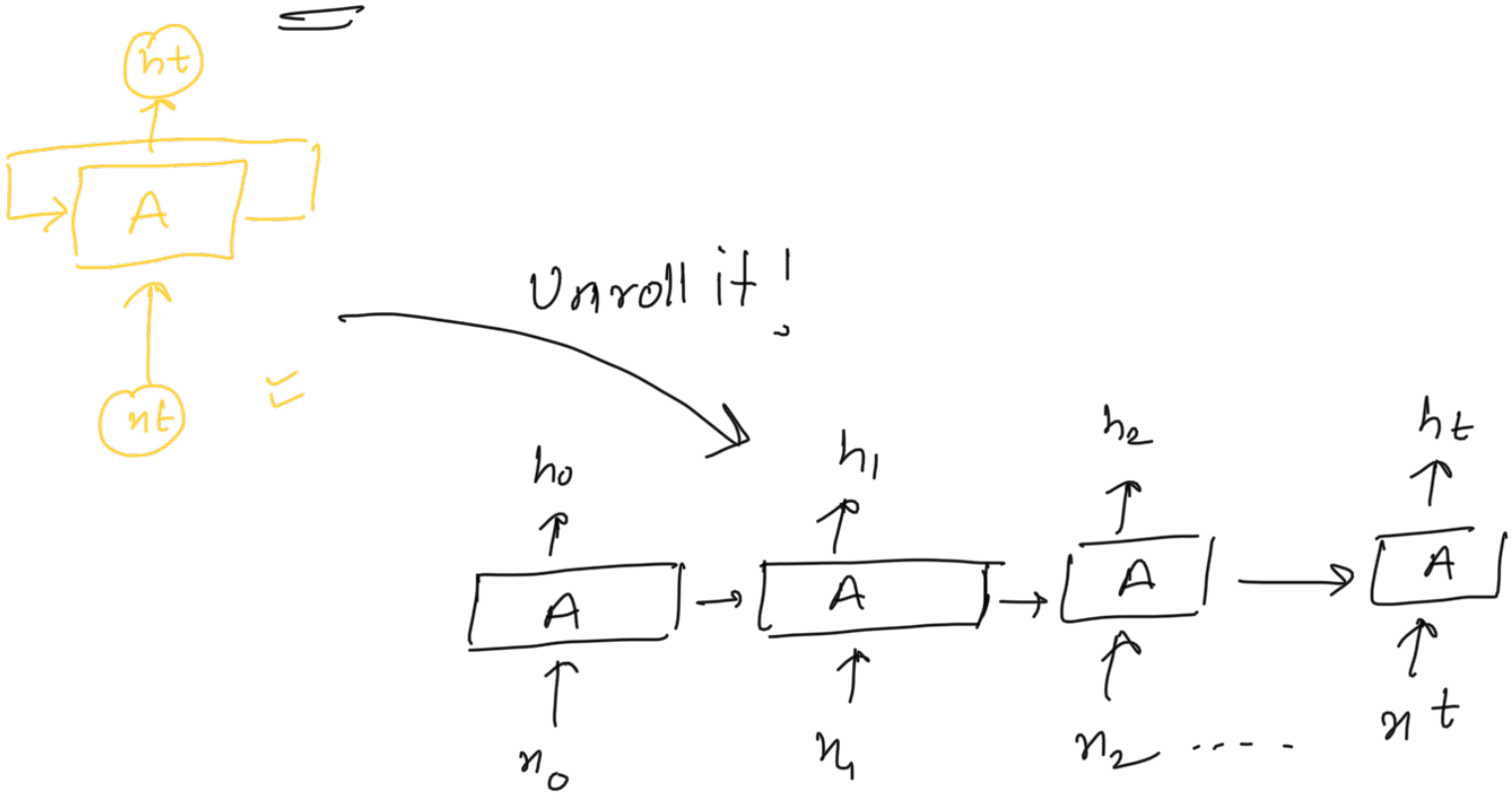
for eg lyrics, etc

↳ It requires us to understand previous input to understand current input

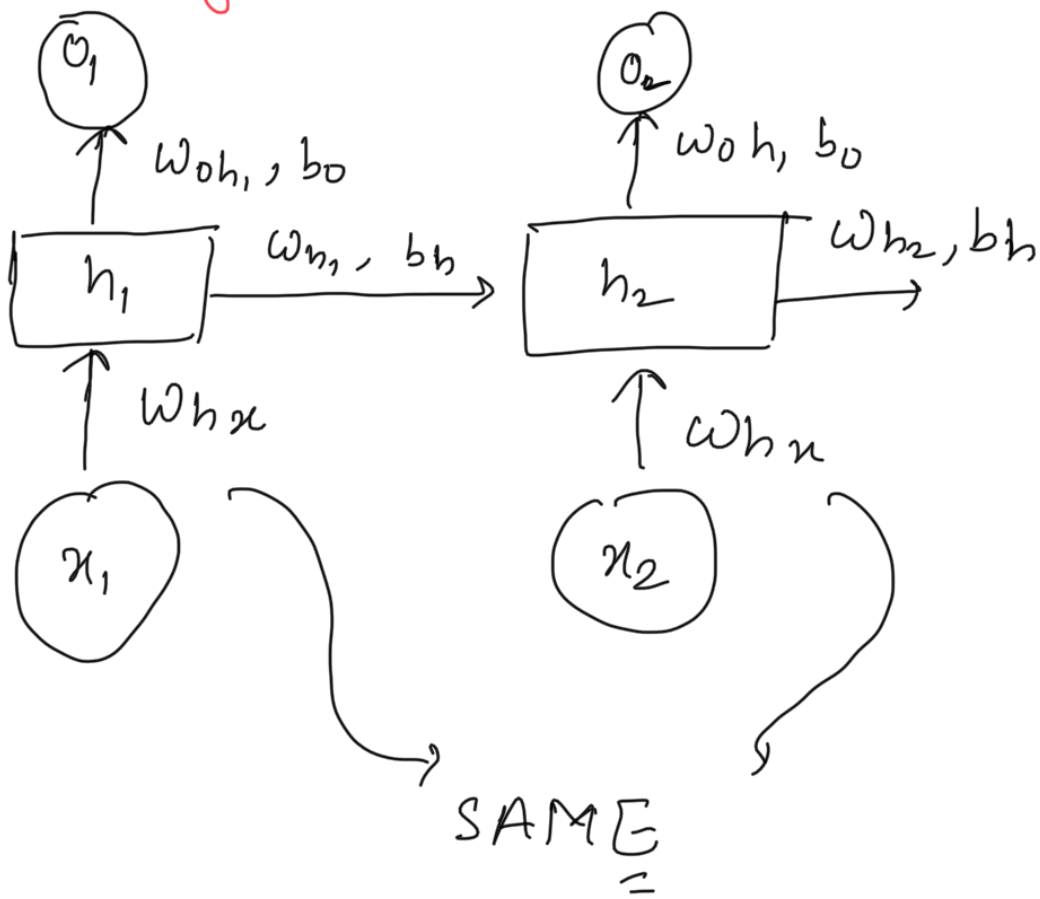
↓ SOLUTION

RNN!

Basic structure



Hidden Layers in the RNN have "SAME WEIGHTS"



So it is clear that

$$h_t = f(h_{t-1}, x_t)$$

("MAGIC!!")

Applying activation function we get

$$h_t = \tanh(W_{hh} h_{t-1} + W_{xh} x_t)$$

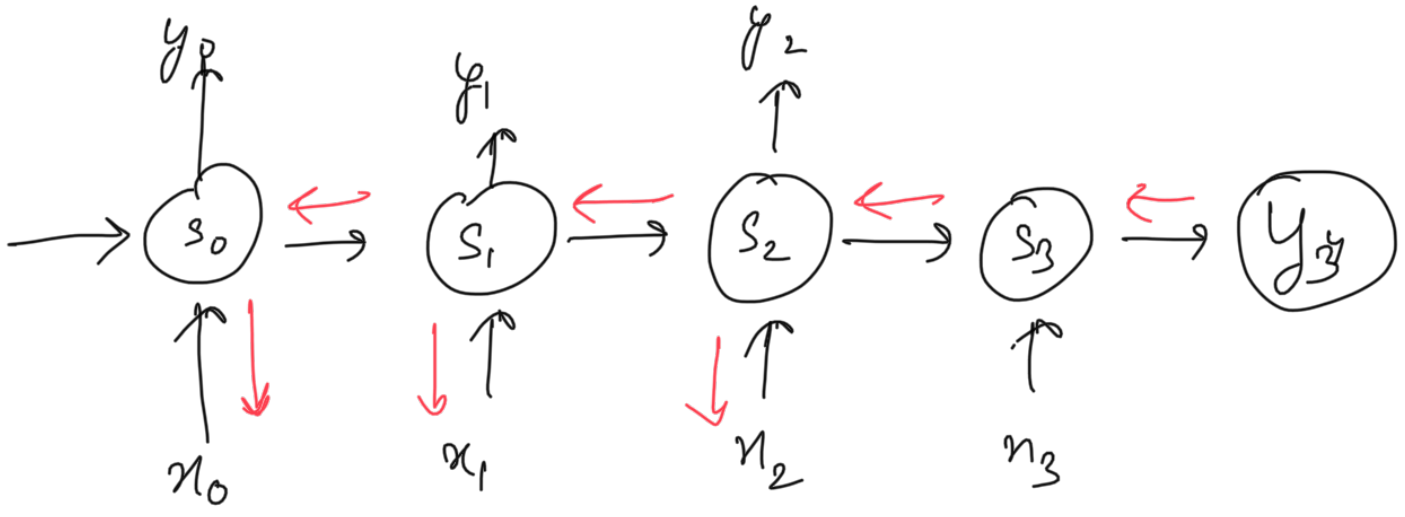
Output

$$y_t = W_{hy} \times h_t$$

see notations from the diagram

Back Propagation Through Time

Backprop for RNN!



Steps :-

1) Forward propagate to get E_3

2) Calculate the error $e_3 = y_3(\text{desired}) - y_3$

3) $\Delta w^3 = \eta \underbrace{\left(\gamma(3)(1-\gamma(3))\, e_3 \right)}_{\text{= differential term}} \times x_2$

4) $\Delta a^3 = \eta \left(\gamma(3)(1-\gamma(3))\, e_3 \right) \times y_2$

using this update

$\simeq w^2, g^2, w^1, g^1$

RTRL