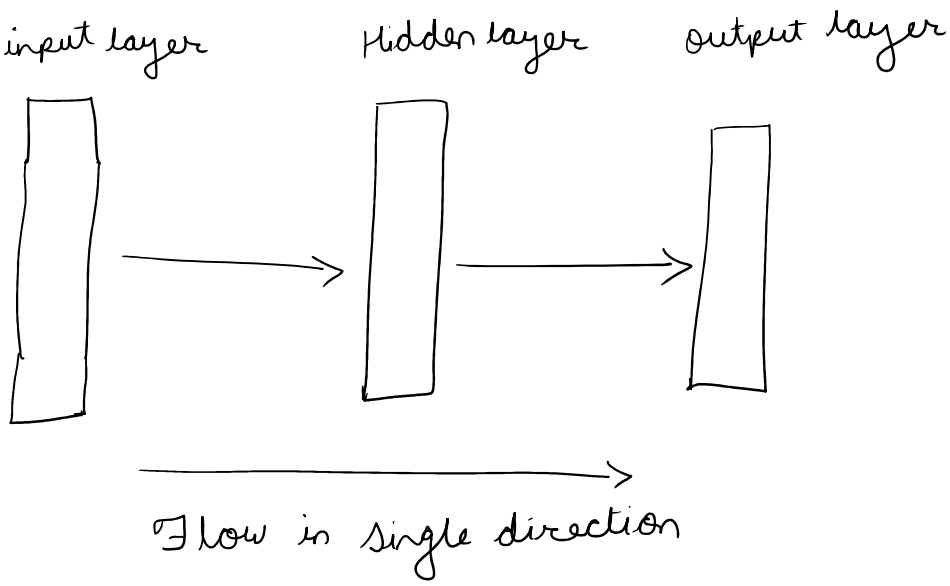
22 January 2024 19:29

RNN → Recurrent Neural Network

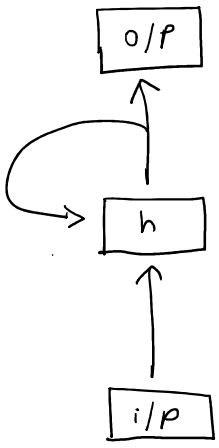
Why RNN?

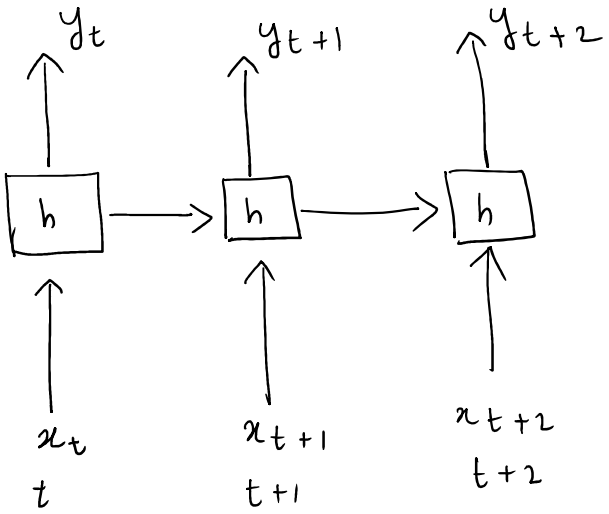
1. To handle sequential data.
2. It considers previous hidden state + current state

Simple feed forward neural network



RNN :

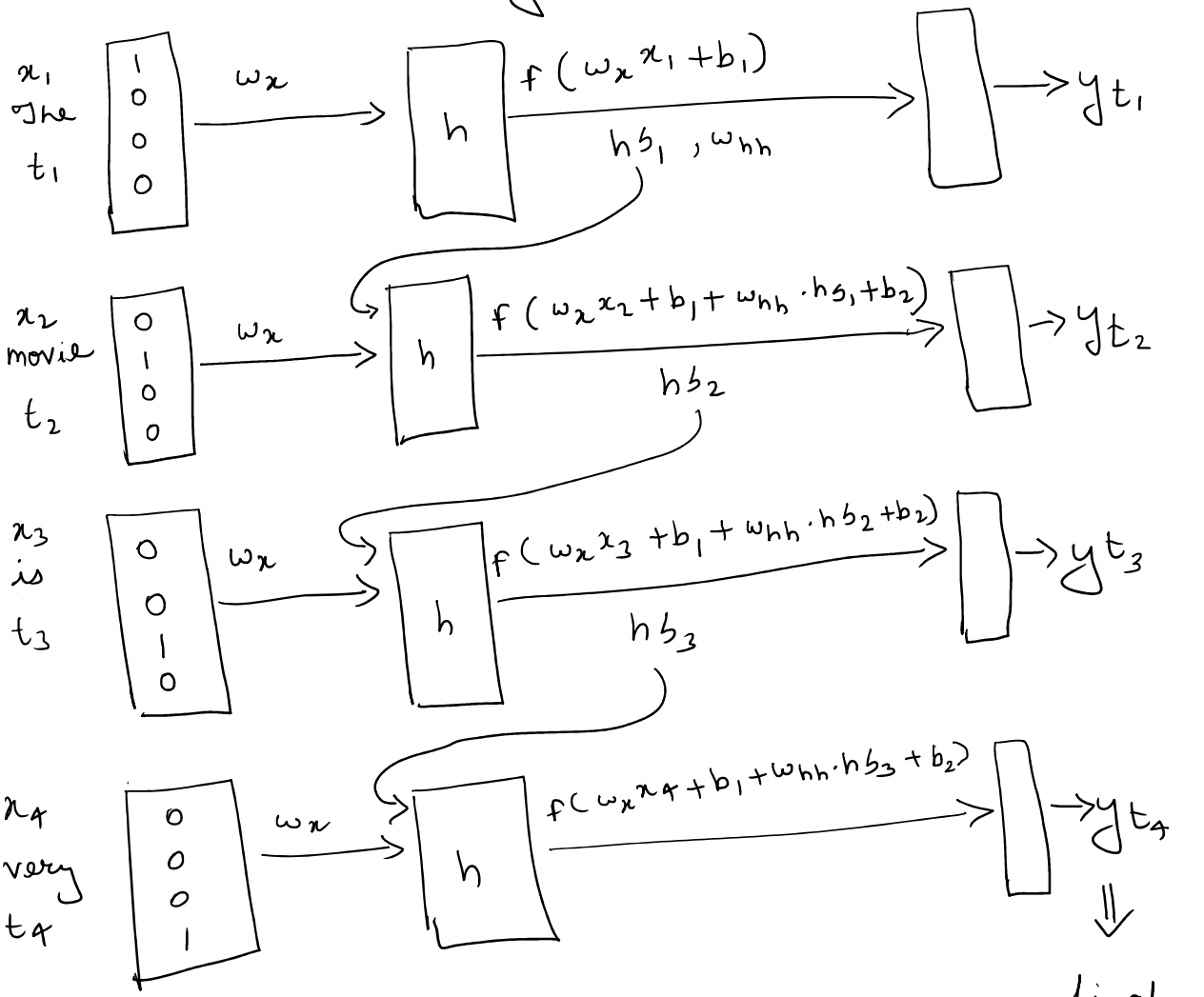




$x \rightarrow w_x, b_1$

$hs \rightarrow w_{hh}, b_2$

The movie is very - - good - -



⇓

final prediction

Types of RNN :: →

① One to one RNN (Vanilla RNN)

② One to many RNN

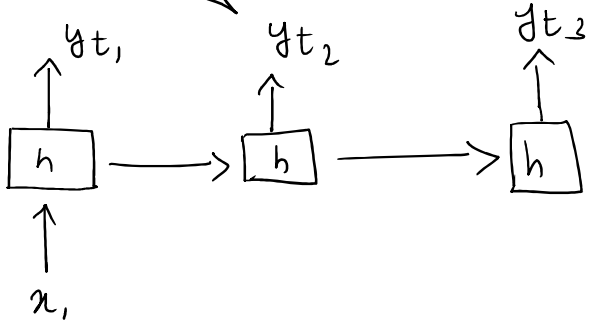
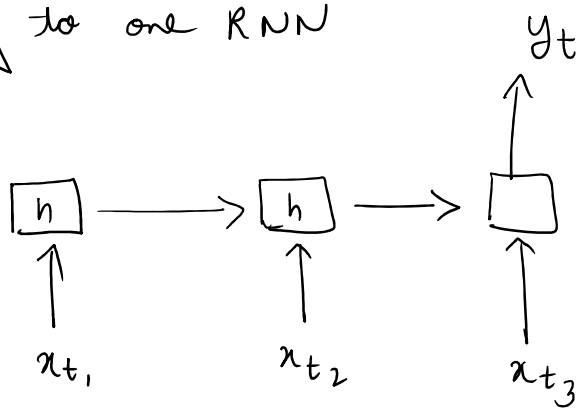


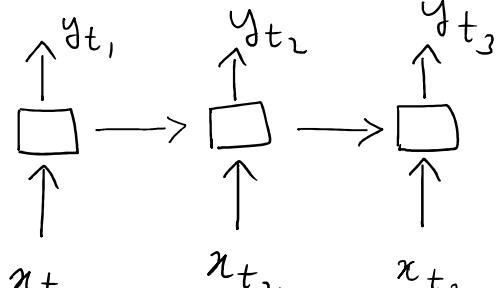
Image captioning

③ Many to one RNN



Sentiment Analysis

④ Many to many RNN

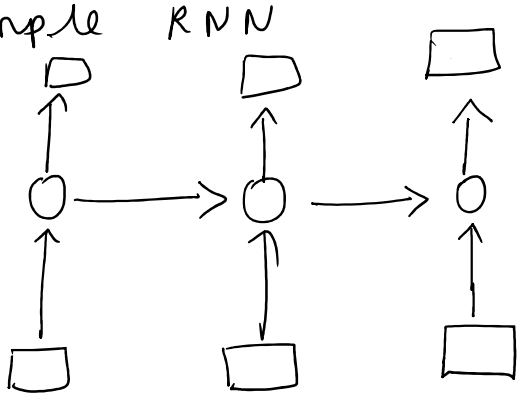


Language translation

x_{t_1} x_{t_2} x_{t_3}

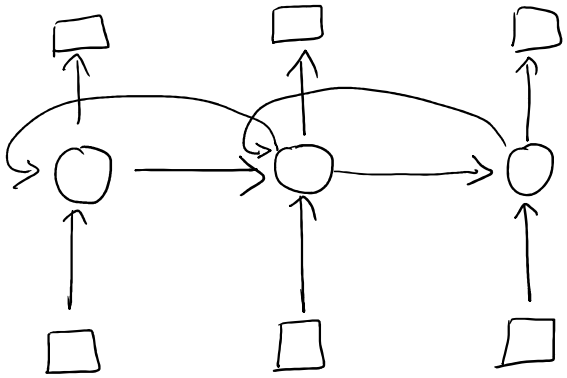
Bidirectional RNN :->

Simple RNN



Rajesh is good in studies. He will be successful.

Rajesh is good in studies. He will be successful.



Drawbacks of RNN:->

① Vanishing gradient.

learning ↓ learning 0

learning ↓ learning 0

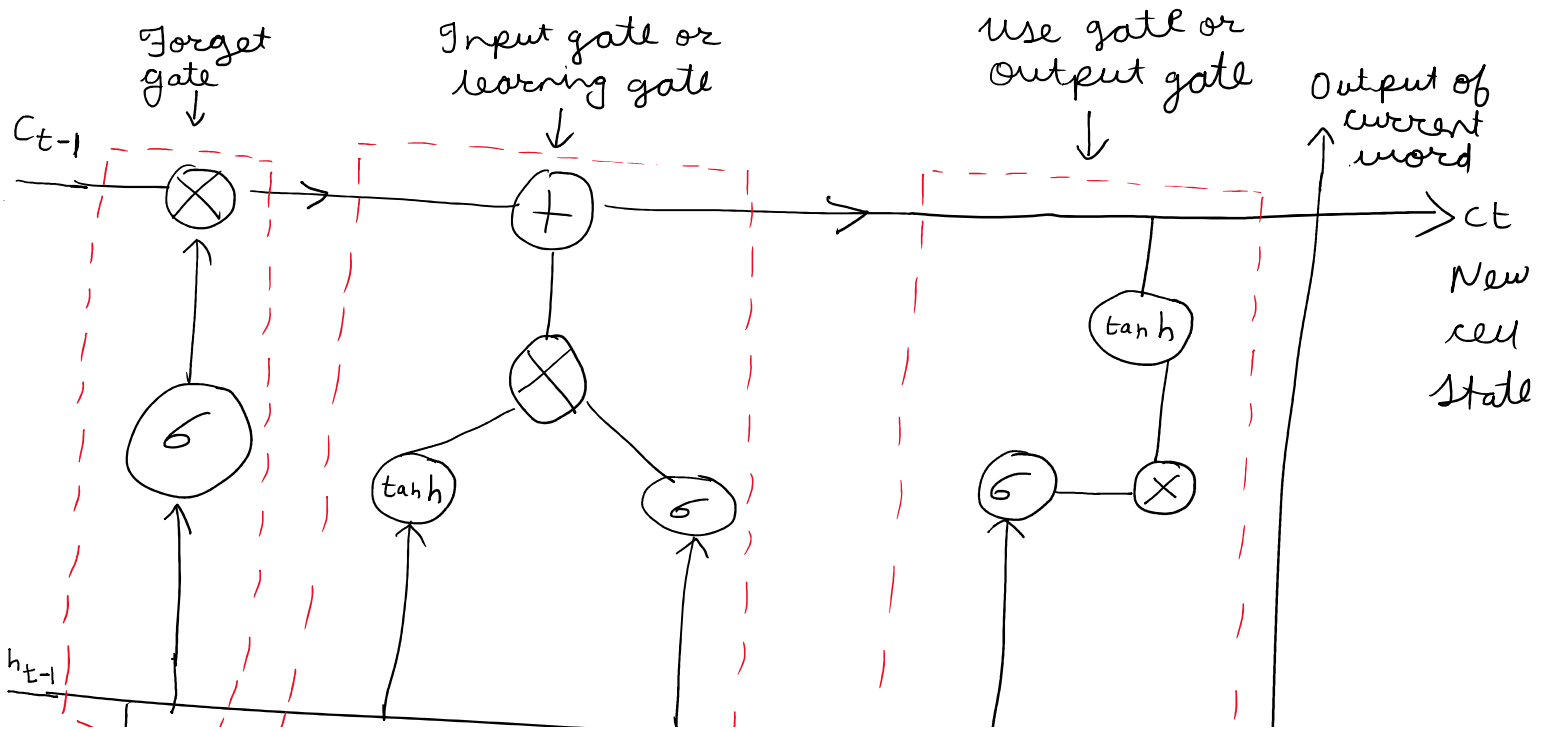
② exploding gradient

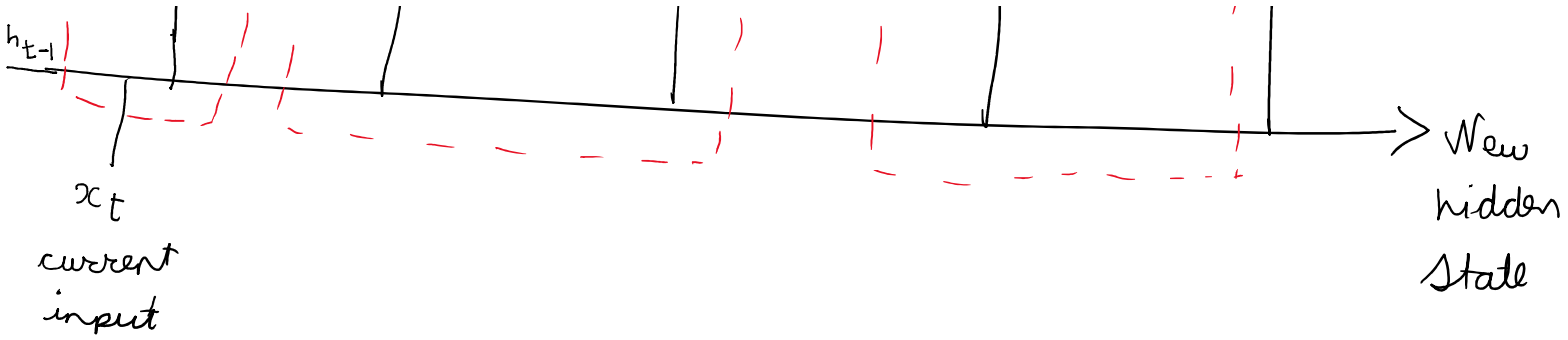
LSTM ⟶ Long Short Term Memory

Cell state ⟶ long term memory

Johny is really good student he is good in maths. He will be successful in future.

// Johny is really good student he is good in maths. He will be successful in future

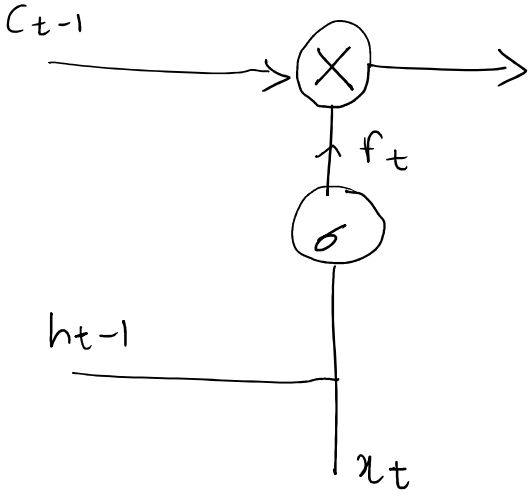




where C_{t-1} => Previous cell state

h_{t-1} => Previous hidden state

① Forget gate



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

$$f_t = [0, 1]$$

Rajesh likes Maths but his friend likes History

⇓

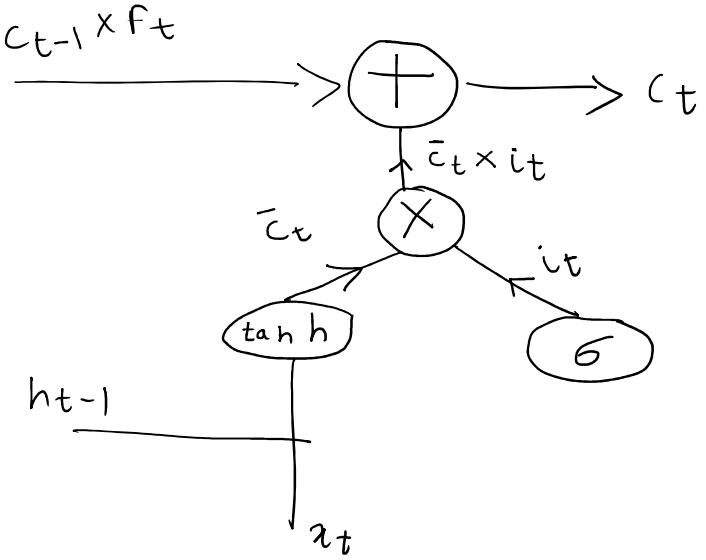
Rajesh likes Maths but he doesn't like History

⇓

Memory cell => 

Add / Remove information

Input gate : Creates new cell state



$$\bar{C}_t = \tanh\left[w_c[h_{t-1}, x_t] + b_c\right]$$

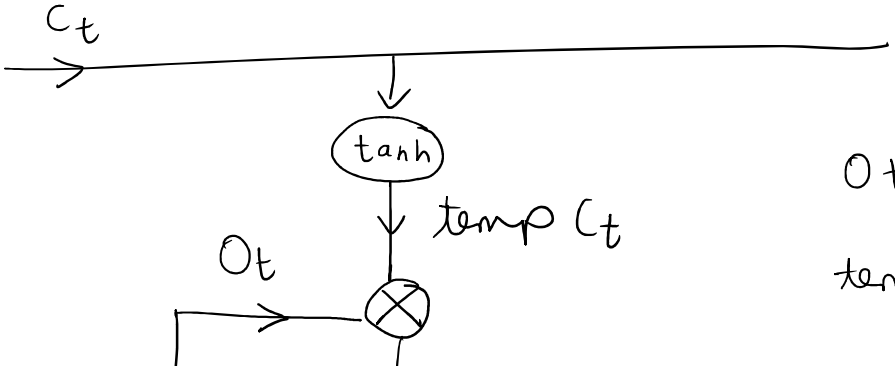
$$i_t = \sigma\left[w_i[h_{t-1}, x_t] + b_i\right]$$

$$\bar{C}_t = [-1, +1]$$

New cell state $\rightarrow C_t = \bar{C}_t \times i_t + C_{t-1} \times f_t$

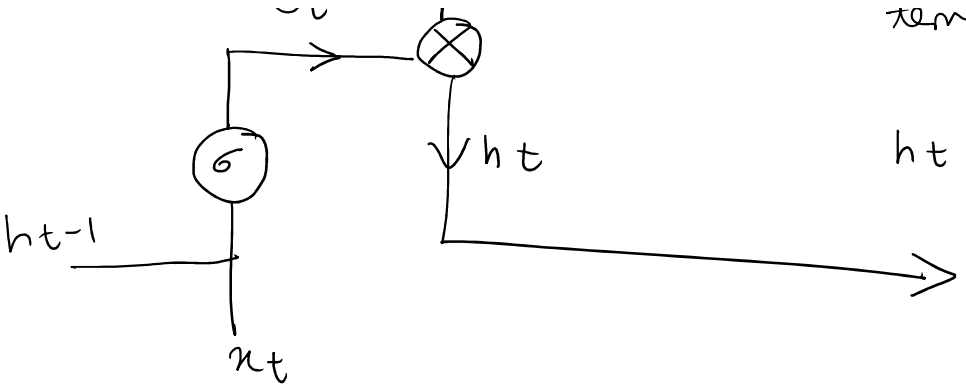
Rajesh likes Maths but his friend likes History

(3) Output gate => New hidden state



$$O_t = \sigma\left[w_o[h_{t-1}, x_t] + b_o\right]$$

$$\text{temp } C_t = \tanh(C_t)$$



temp c_t = tanh(c_t)

$h_t = O_t \times temp\ c_t$

Rajesh likes Maths but his friend likes History.

Drawbacks of LSTM:

1. It will take longer time to train.
2. It requires more memory to train.
3. It can get overfit.

- Text classification
- Document classification
- Topic modelling
- NER

Sentiment Analysis + Root cause Analysis