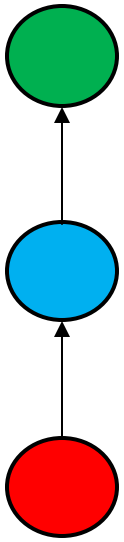
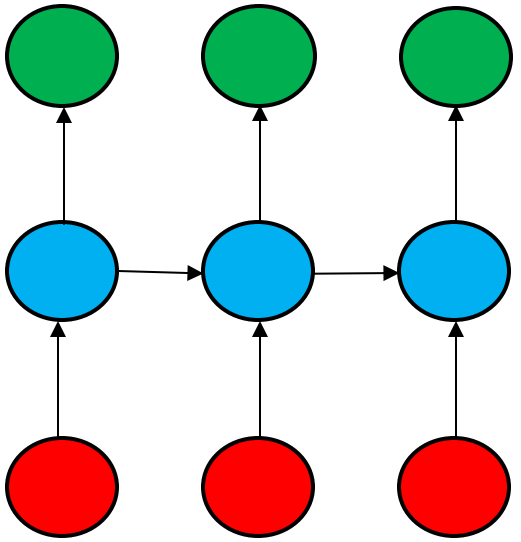
Recurrent Neural Network (RNN) and Long Short Time Memory (LSTM)

BY CECCOTTI ROMAIN, DENELE LUCAS, EL KHIATI AMINE, KUCUKAL BURAK AND XU LU

RNN

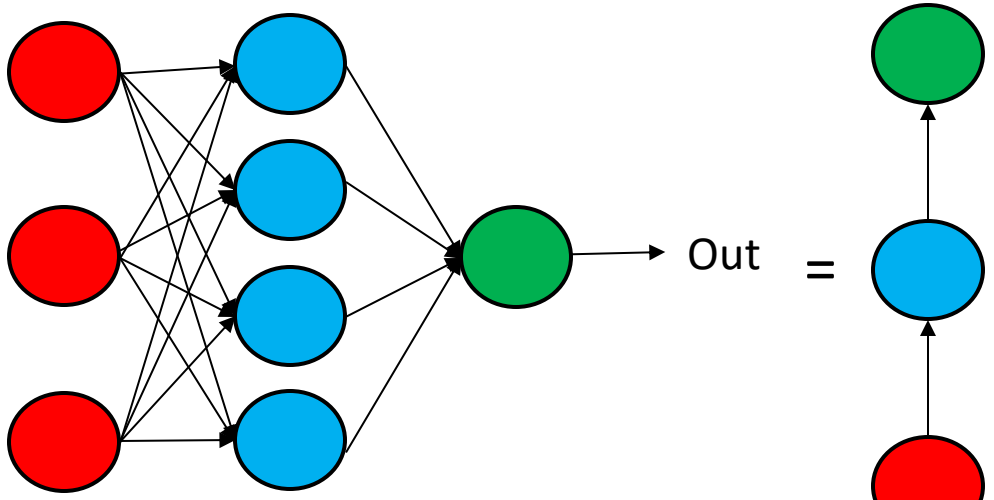


One to One

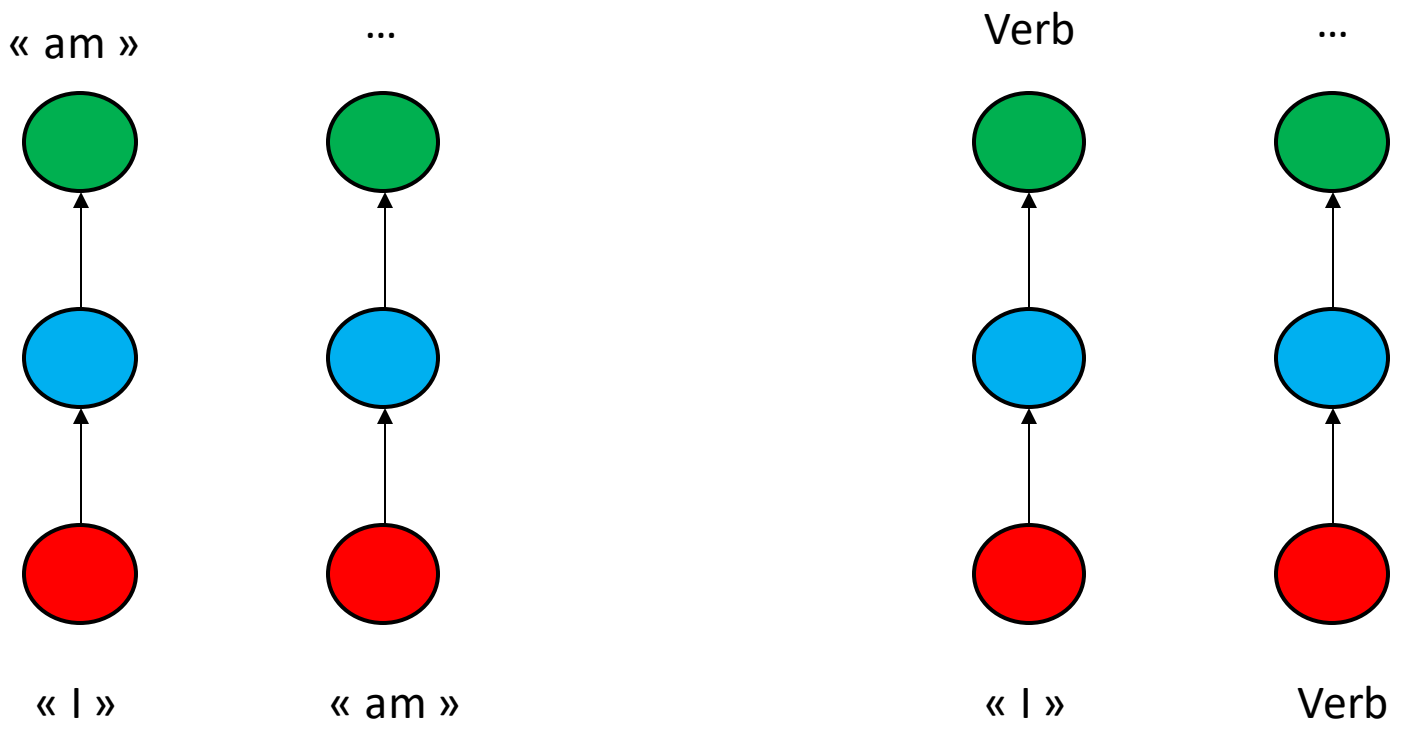


Many to Many

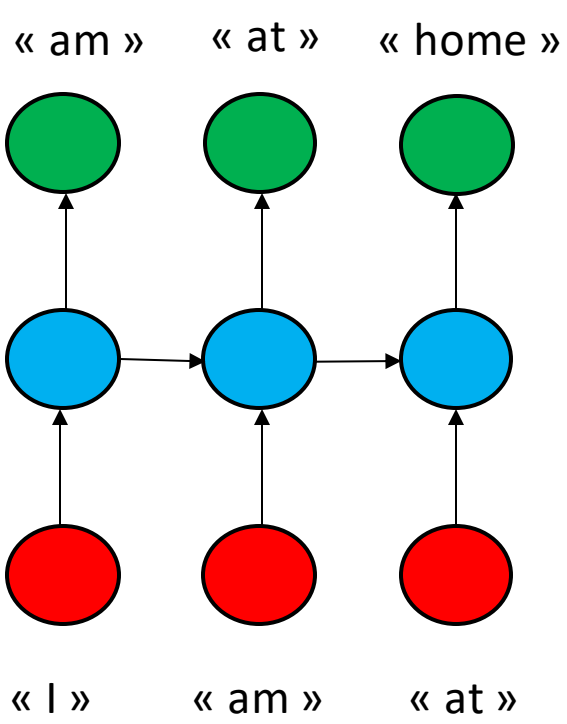
RNN : The feedforward

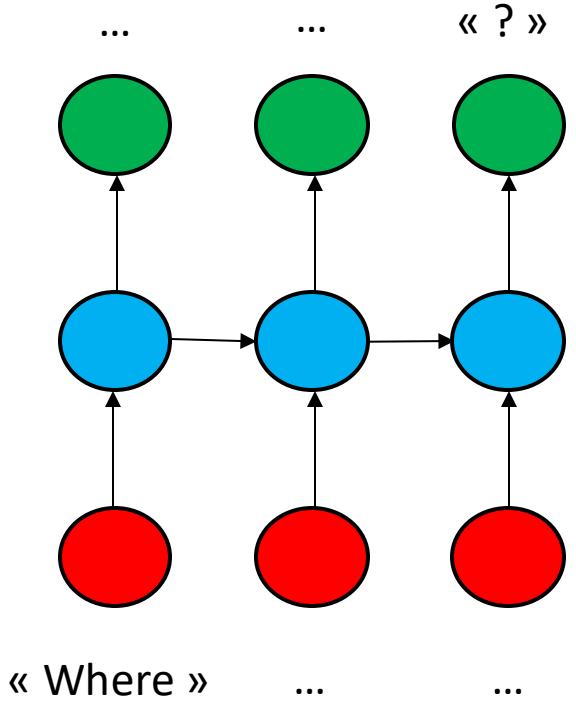


RNN : The feedforward

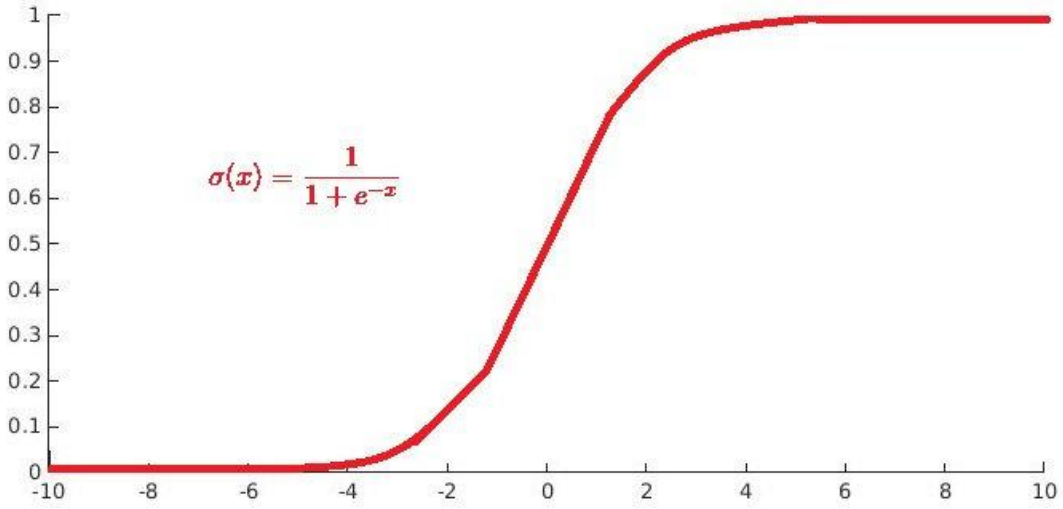


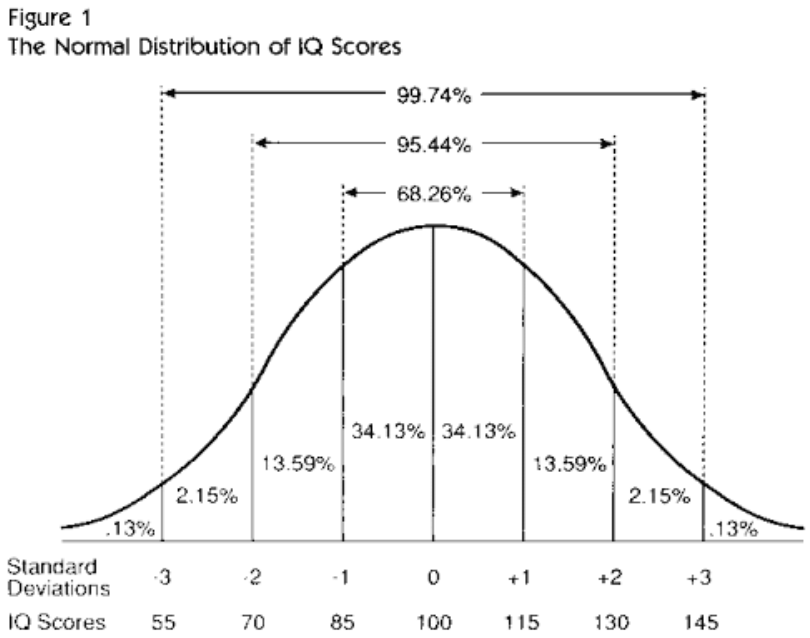
RNN : The recurrent neural network





RNN : Vanish gradient





LSTM : Three main operations

- Forget Gate
- Input Gate
- Output Gate

LSTM : The forget Gate

- Gets t-1 output value
- Apply a sigmoïd function to it
- Saves the result in the memory called « C »

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f)$$

LSTM : Input Gate

- Aplly tanh and adds a bayes to the memory
- Proposes new interesting values
- Aplly the input gate function to these values

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

LSTM : Output Gate

- Apply the sigmoïd function on the memory
- Apply tanh on the actual memory
- Multiply the previous results

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$