Introduction to the basics of AI - S8
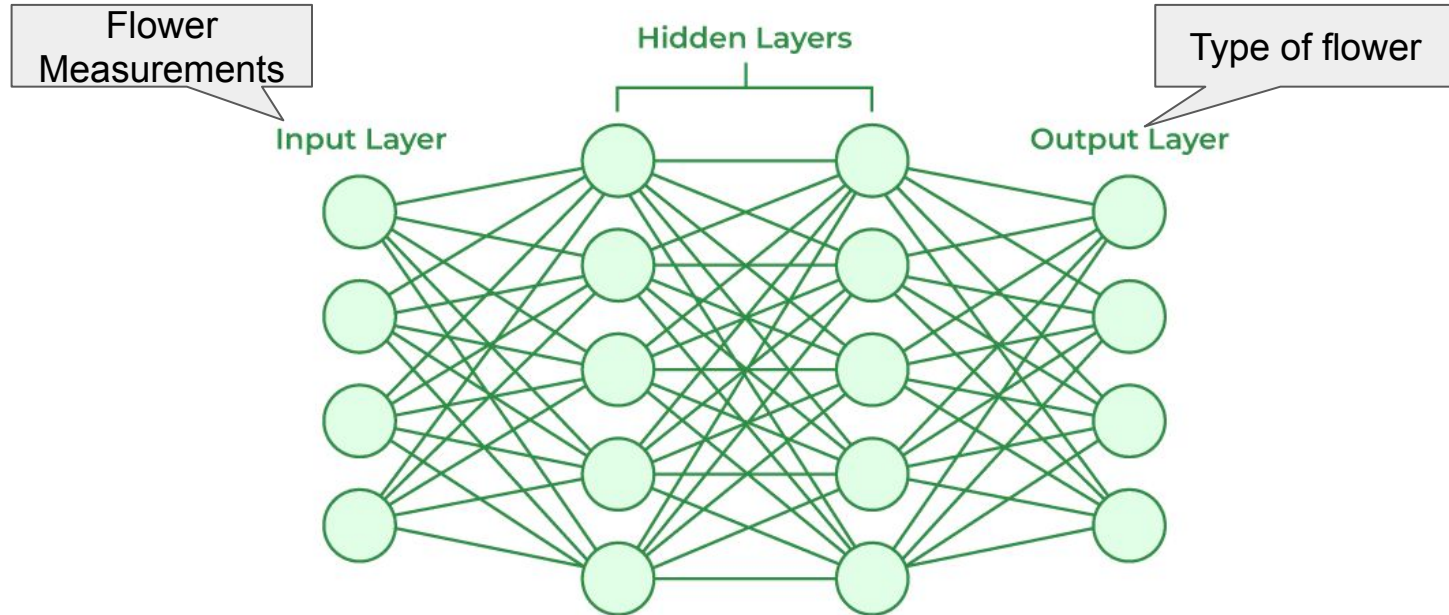
Z. TAIA-ALAOUI

Outline for today's course

- Non-Sequential Vs Sequential Data

- Classical RNNs

- LSTMs

- Implementation

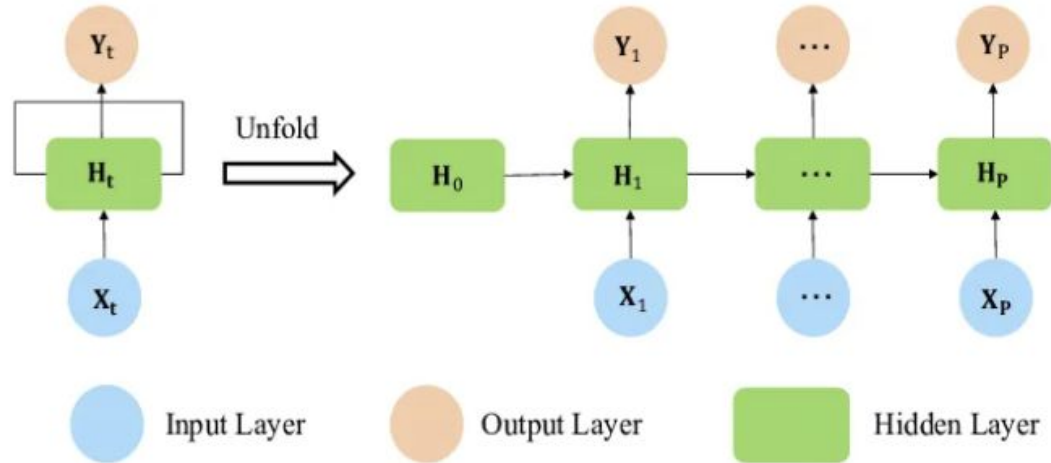# Fully Connected Networks (Non-Sequential Data)

- There is no time involved in x and y
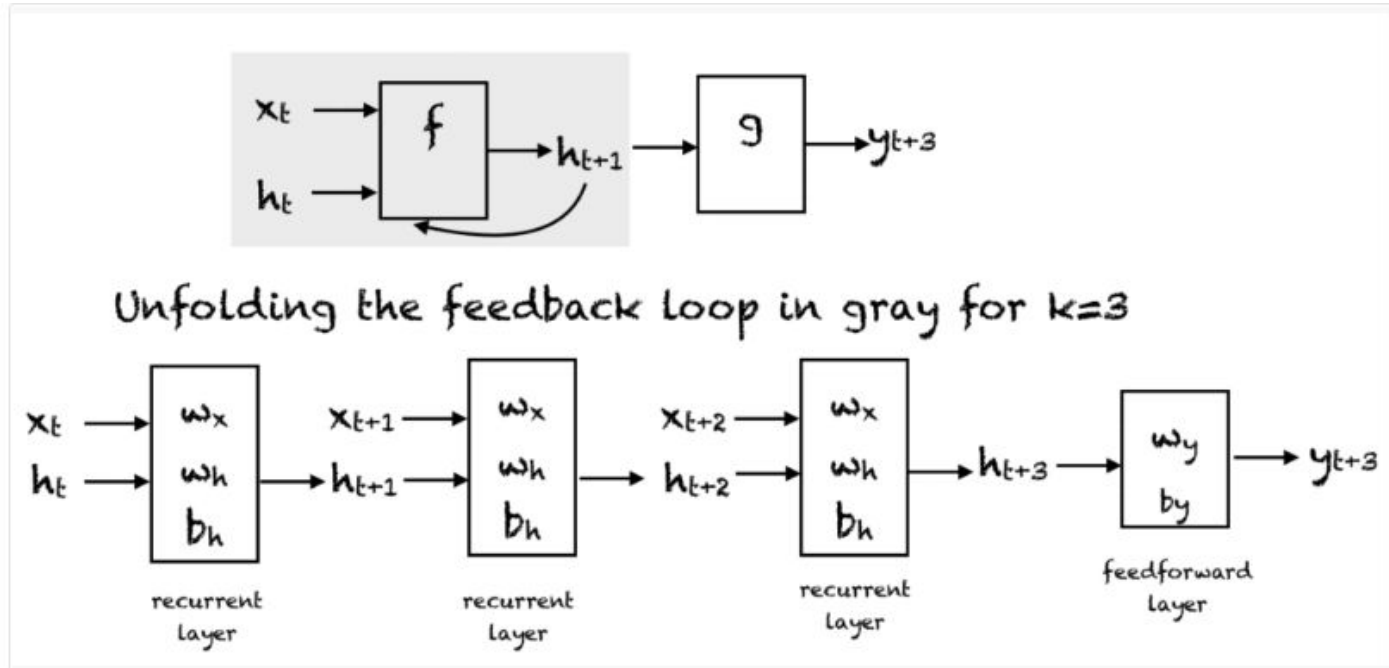
$$\hat{y} = \sigma(w^T x + b)$$

Unfolding of an RNN Cell

- In RNNs we usually talk about cells rather than unit neurons
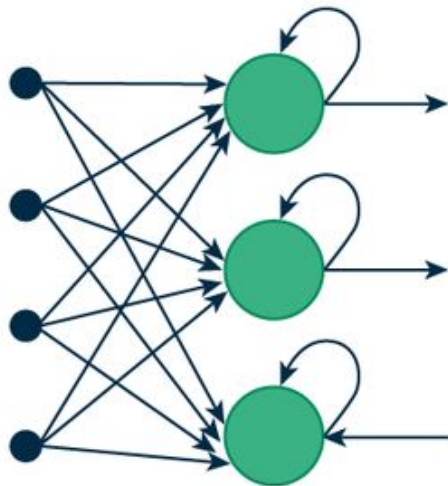- Each cell can be deployed in time such that:
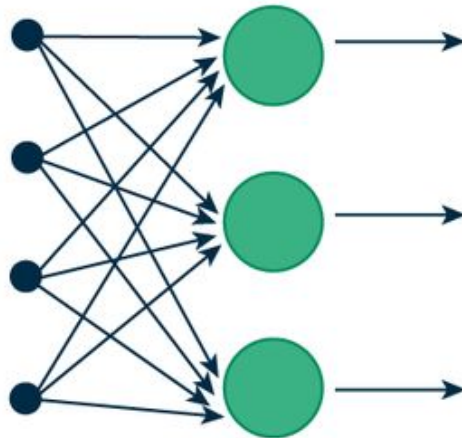
Unfolding of an RNN Cell

Recurrent Neural Networks (Sequential Data)

$$p(y_t/x_t, h_{t-1})$$
$$p(y_t/x_t, y_{t-1})$$

$$p(y/x)$$

(a) Recurrent Neural Network

(b) Feed-Forward Neural Network

Learning with RNNs

- Forward Pass

$$y_t = g(W_{hy} h_t)$$

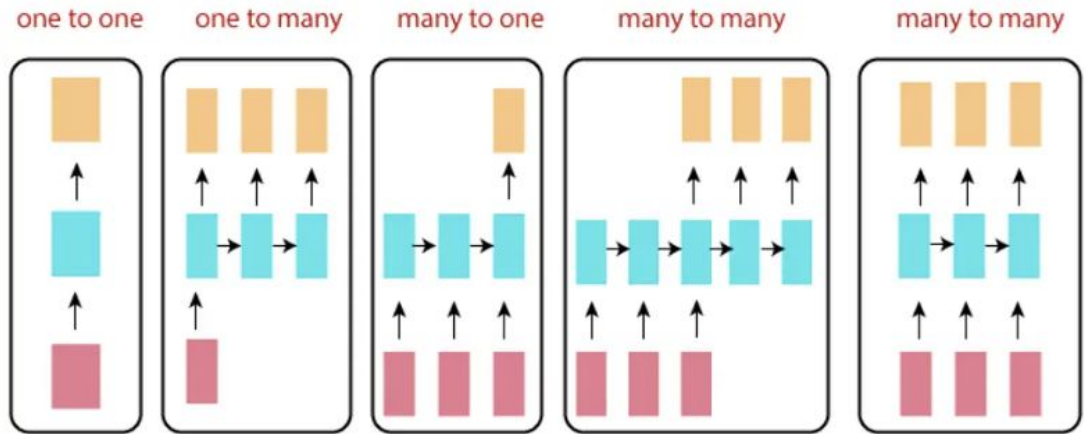$$h_t = f(W_{hh} h_{t-1} + W_{xh} x_t)$$

- Backward Pass

$$L = Sum(cross - entropy)_{over-time}$$

$$L = \sum_t C_t$$

$$\frac{\delta C_t}{\delta w_{ij}} = \frac{\delta C_t}{\delta y_j} \frac{\delta y_j}{\delta h_j} \frac{\delta h_j}{\delta w_{ij}}$$

Applications

- Forecasting Stocks
- Forecasting physiological signals
- Forecasting the next word in a sentence
- Forecasting the weather
- Denoising
- Reconstructing Missing Data

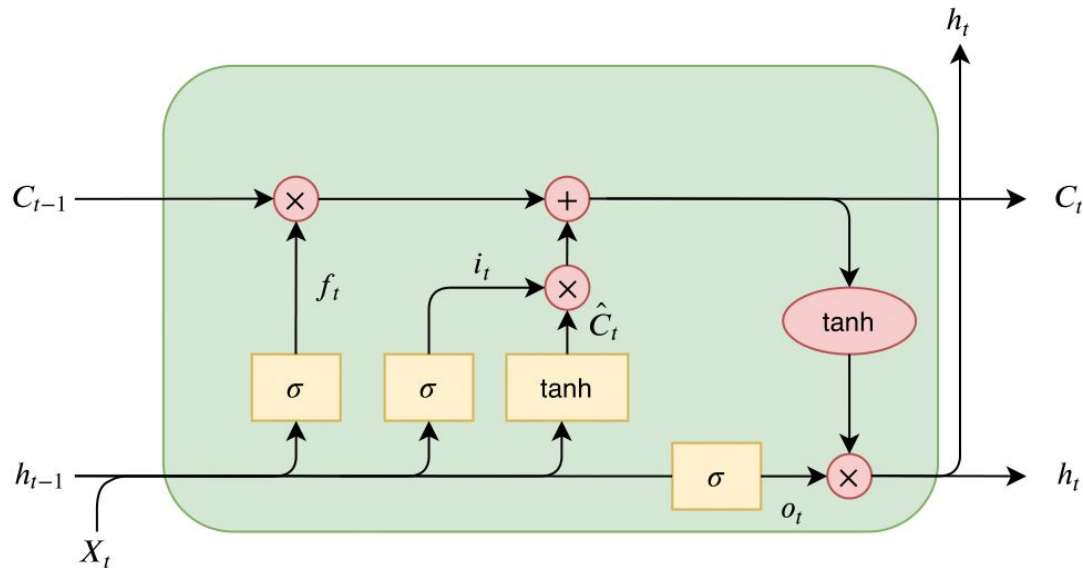| one to one | one to many | many to one | many to many | many to many |

Some limitations of RNNs

- Vanishing or exploding gradients: back propagation through time (BPTT)

- Long-term Memory
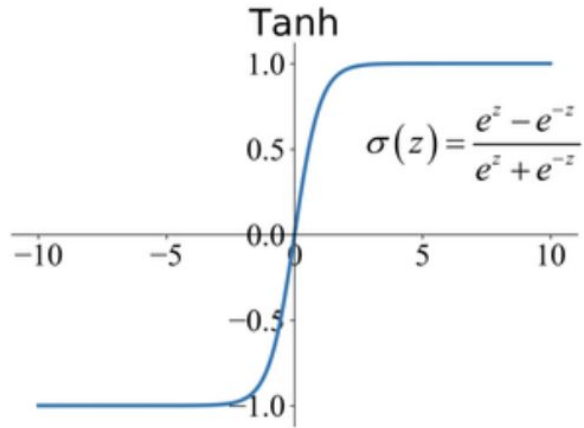
- High Cost due to long BPTT

LSTMs

$$f_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f)$$
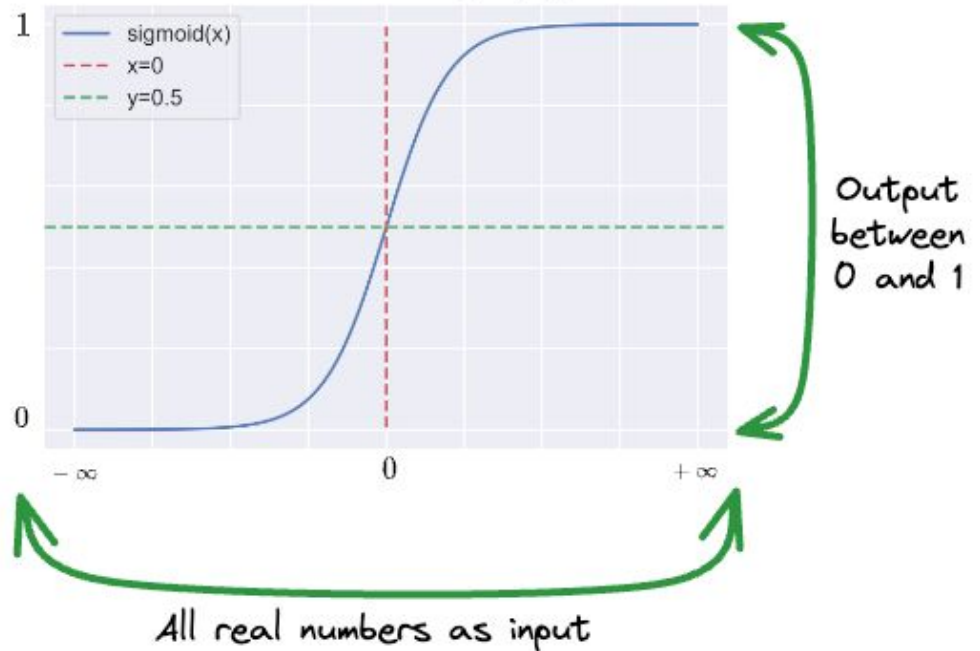
$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o)$$

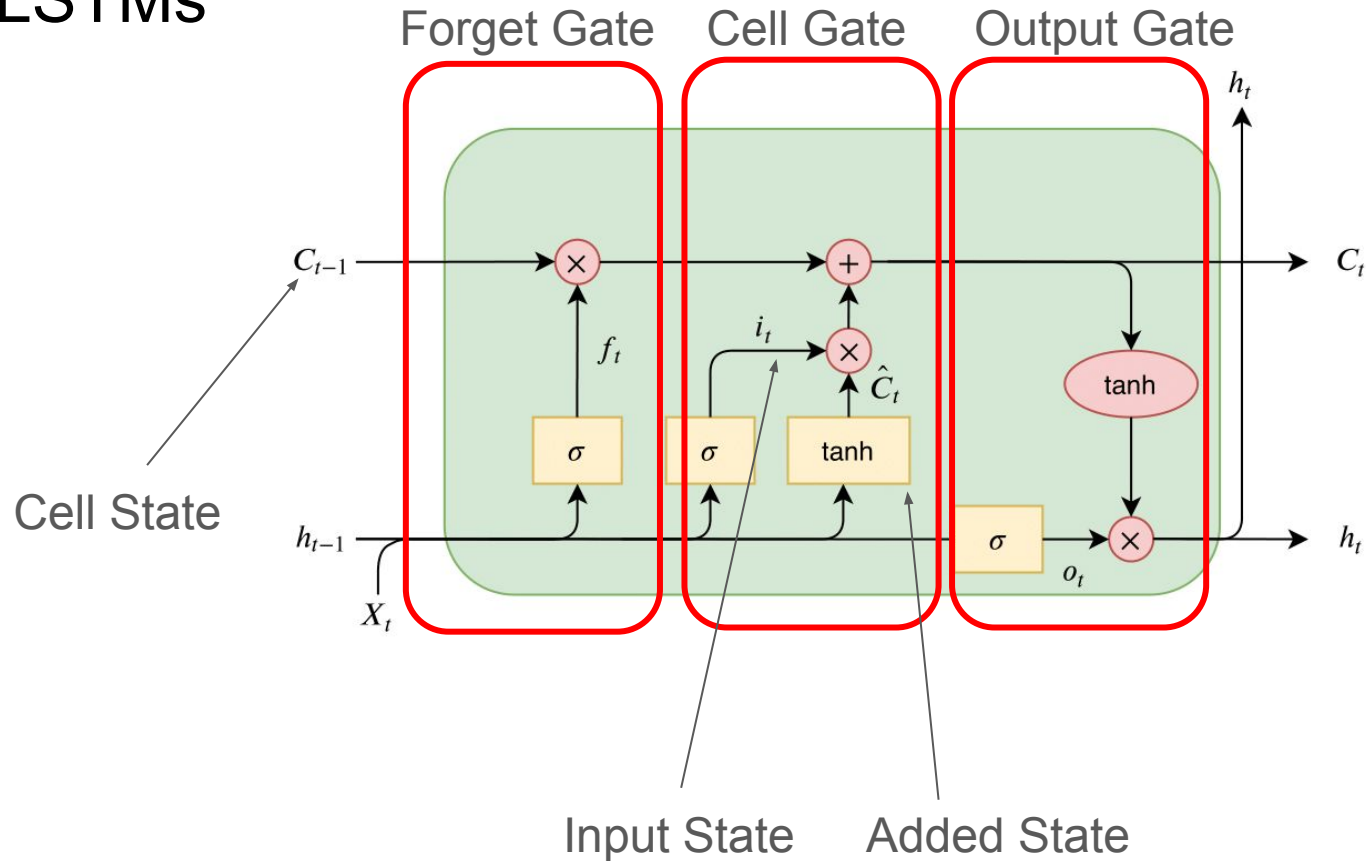$$\hat{C}_t = \tanh(W_C \cdot [h_{t-1}, X_t] + b_C)$$

Activations of an LSTM

Tanh

$$\sigma(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$$

- sigmoid(x)
- x=0
- y=0.5

Output between 0 and 1

All real numbers as input

LSTMs

Tutorials

https://colab.research.google.com/github/kmkarakaya/ML_tutorials/blob/master/LSTM_Understanding_Output_Types.ipynb#scrollTo=9grVun4vXk7n

https://github.com/mapr-demos/predictive-maintenance

https://machinelearningmastery.com/lstm-for-time-series-prediction-in-pytorch/

Resources

https://dotnettutorials.net/lesson/recurrent-neural-network/

https://medium.com/@adachoudhry26/getting-started-with-ai-building-an-rnn-from-scratch-and-practicing-resilience-ba3c10be6a22

https://machinelearningmastery.com/an-introduction-to-recurrent-neural-networks-and-the-math-that-powers-them/

https://colah.github.io/posts/2015-08-Understanding-LSTMs/

https://paperswithcode.com/method/tanh-activation

https://www.dailydoseofds.com/why-do-we-use-sigmoid-in-logistic-regression/