# Long Short Term Memory Model

* Capable model sequential and temporal aspects of data.
* They build up a bigger recurrent neural network(RNN)
* LSTM block is composed of four main components
  1. Cell: remembering values over arbitrary time intervals, “memory” in LSTM
  2. Input gate
  3. Output gate
  4. Forget gate
* Each of three gates are “conventional” artificial neuron, they are the connection between the gates and cells
* An LSTM is well-suited to classify, process and predict time series given time lags of unknown size and duration between important events.
* LSTM widely used for text, videos, and time-series data.
* Basic use of LSTM network:
  1. Language modeling
  2. Machine Translation
  3. Image captioning
  4. Handwriting detection
  5. Image Generation
  6. Automatic question answering
  7. Video to text extraction
* Two major issue that LSTM solve for Recurrent Neural Networks
  1. RNN suffers from gradient problem, it limits RNN effectiveness when it needs to go back deep in the context
  2. No control of what part of the context need to be carried forward and what part of the context needs to be forgotten

P.s Vanishing gradient problem: is a difficulty found in training artificial neural networks with gradient-based learning methods and backpropagation

* By using gated units would solve the two issues, LSTM also carry some good feature a RNN has such as the use of backpropagation through time, network architecture.

# Recurrent neural network

* It’s a class of artificial neural network, connect unit from direct cycle
  + Direct cycle: a close walk with no repetitions of vertices and edges allowed other than the repetition of the starting and ending vertex or as the set of edges in such a walk
* RNN can use their internal memory to process arbitrary sequences of inputs. They are applicable to talks such as unsegmented connected handwriting recognition or speech recognition
* There are different types of architectures: Fully recurrent, Recursive, Hopfield, Bidirectional associative memory, Elman networks and Jordan networks

# Feedforward neural network

* It is an artificial neural network wherein connections between the units do not form a cycle.
* The information only move in one direction, forward, from the input nodes through the hidden nodes and to the output nodes, no cycle or loop.
* Single-layer perceptron
  + Consist of a single layer of output nodes, input are fed directly to the outputs via a series of weights
* Multi-layer perceptron
  + It can compute a continuous output instead of a step function