LSTM

* Designed to overcome limitations of RNN:
  + Like vanishing and exploding gradient
  + Complex training in the RNN
  + Difficulty of processing long sequences

Diagram

Description automatically generated

* In the LSTM setting we have:
  + Tanh = -1 to 1
  + Cell state: upper bar is the cell state (information flows though this path)
  + Forget gate: (let something in or not) – helps us to find out what to forget and what to let ind
  + Input gate: controls the information that gets though
  + Output gate: controls what information that do enter

Diagram

Description automatically generated

* Each of the gates can either be zero or one: meaning that as the information pases though the upper path it can either be multiplied by the gate if it is equal to one or nothing happens to the information if the gates are equal to zero.

Diagram

Description automatically generated

Diagram

Description automatically generated

For information:

* https://www.youtube.com/watch?v=Mubj\_fqiAv8&list=PLeo1K3hjS3uu7CxAacxVndI4bE\_o3BDtO