**Brief Overview of LSTMs**

**Simplified Version**

**Problem with this version:** Does not accurately differentiate between hidden state and outputs (dotted line shown below – more on this at the bottom)

A screenshot of text

Description automatically generated

**Actual Version**

![A picture containing clock

Description automatically generated]()

**How these two pictures interact with each other**

**A screenshot of a map

Description automatically generated**

**Short Crash Course on LSTMs**

![A picture containing clock

Description automatically generated]()

* Composed of 3 gates:
  + Input
    - Contains a sigmoid function which decides which values to let through (selection of relevant bits)
    - Contains tanh function which decides level of importance of passing material
  + Forget
    - Discover which details need to be forgotten using the sigmoid function
    - Looks at previous state and current input
  + Output
    - Uses the memory of the block (obtained from the forget gate) and input
    - It essentially collects all of the **COLLECTED** (filtered possibilities + forget gate which uses memory) **POSSIBILITIES** (input gate)
    - Has a sigmoid functions which decides which values to let through (0,1) and a tanh function to assign level of importance (-1, 1)
  + What is the difference between the hidden state and the output?
    - The hidden state is the vector representing the set of **collected possibilities** we talked about above
    - Before the output gate applies the sigmoid and tanh function – mathematically that’s what the hidden state is
    - Conceptually