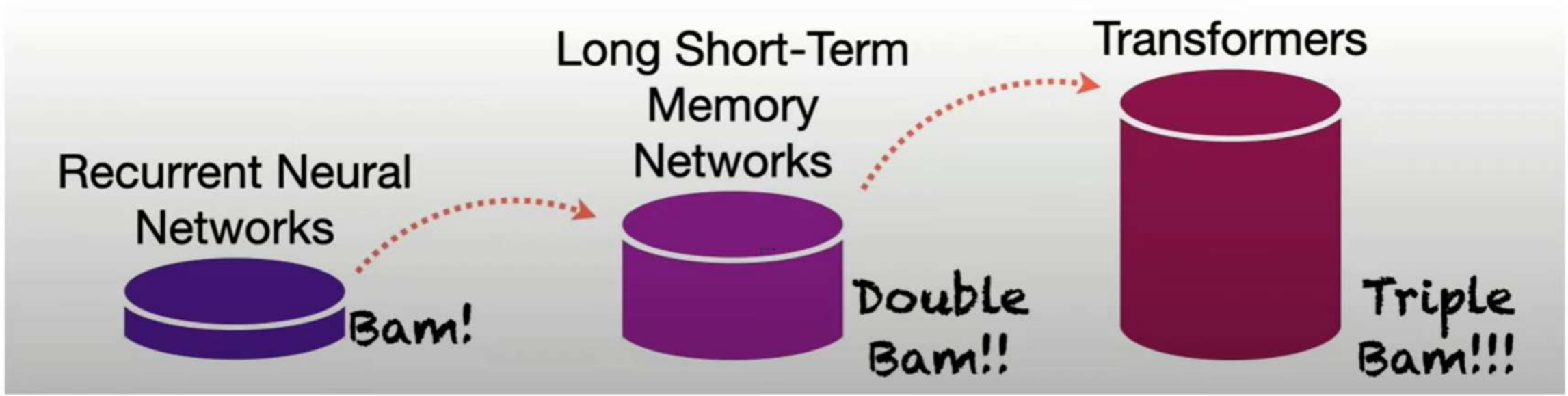
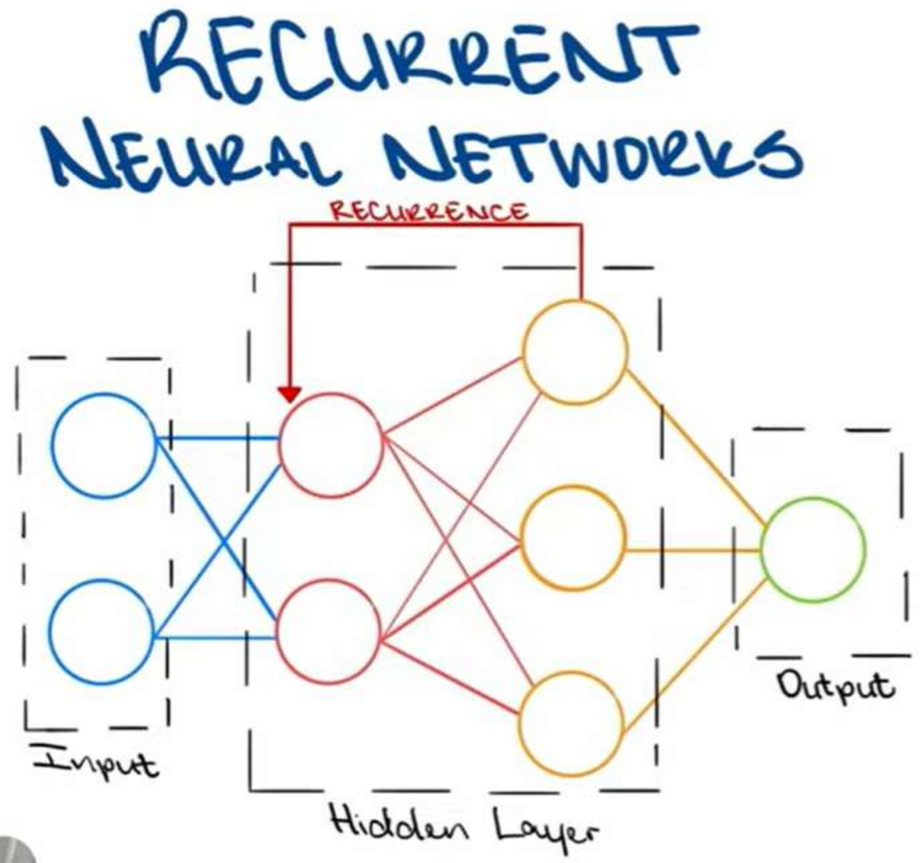

Recurrent Neural Networks
Bam!
Long Short-Term Memory Networks
Double Bam!!
Transformers
Triple Bam!!!
1980
1997
2017

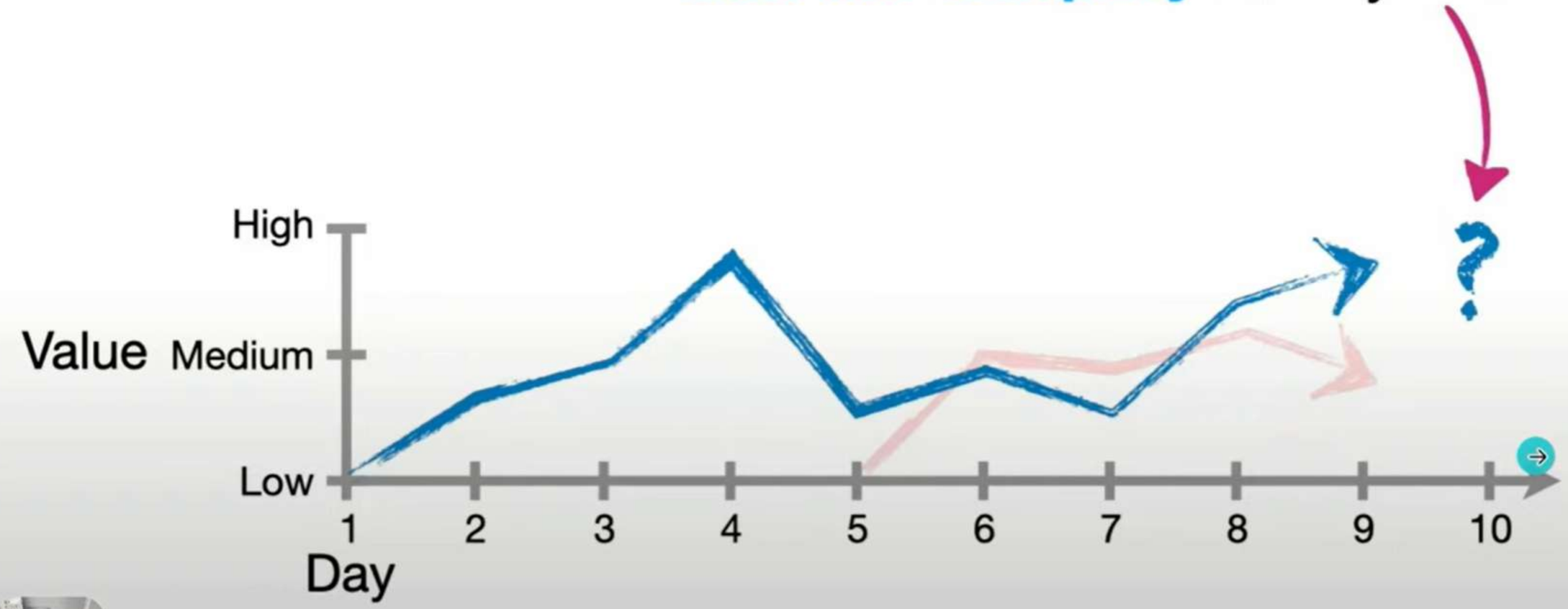
Generative AI History: Recurrent neural networks - 1980 (Memory)



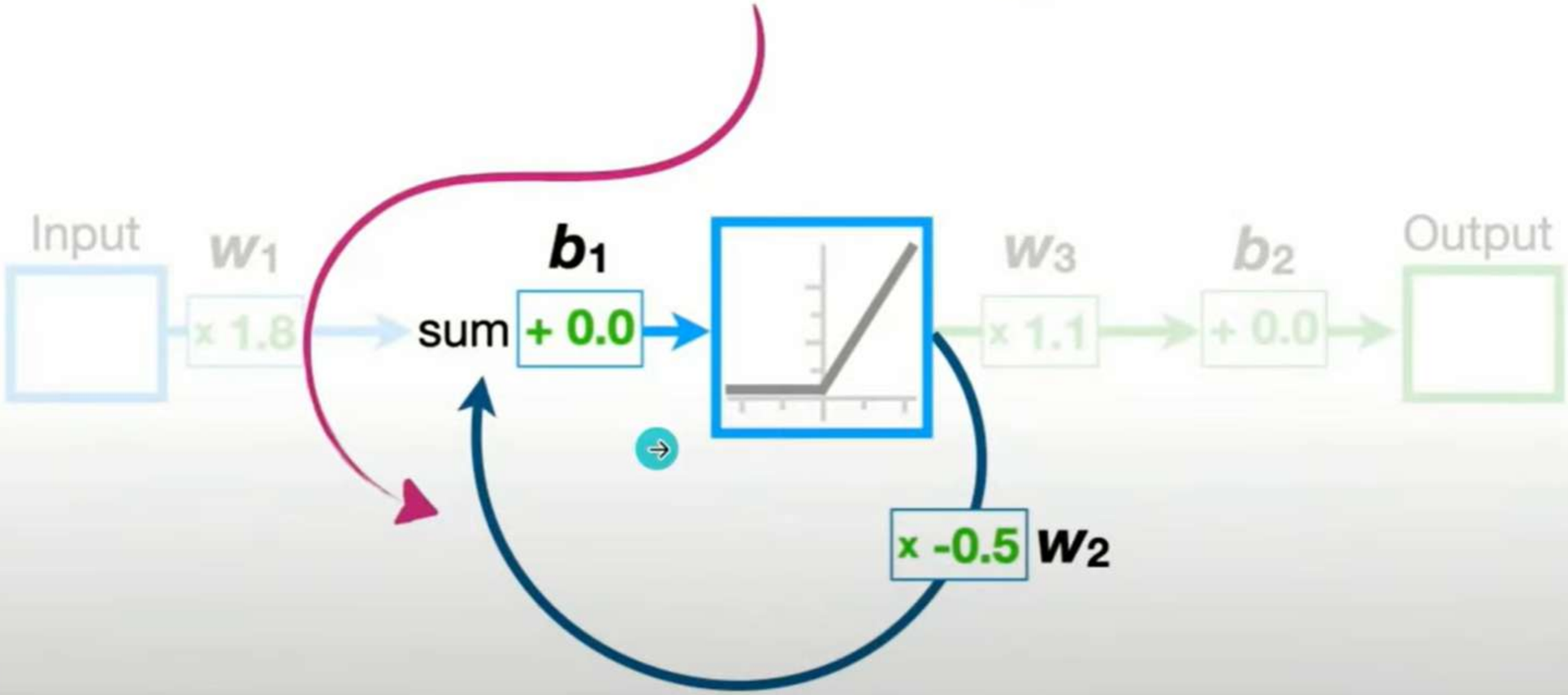
RNNs were designed to work with sequential data by maintaining a hidden state that captures information about previous inputs.

They were a significant advancement for generative models, allowing for the generation of sequences of text, music, and more.

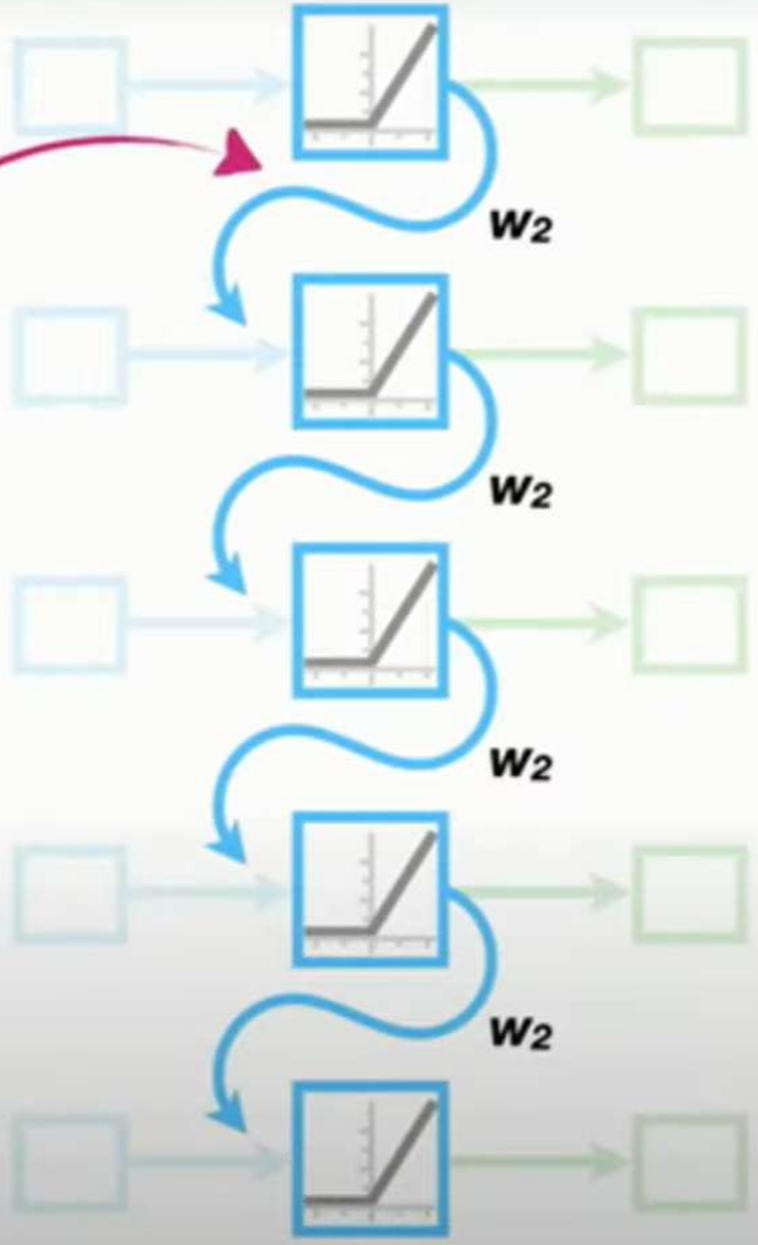
In other words, if we want to predict the stock price for the **blue line company** on day **10**...



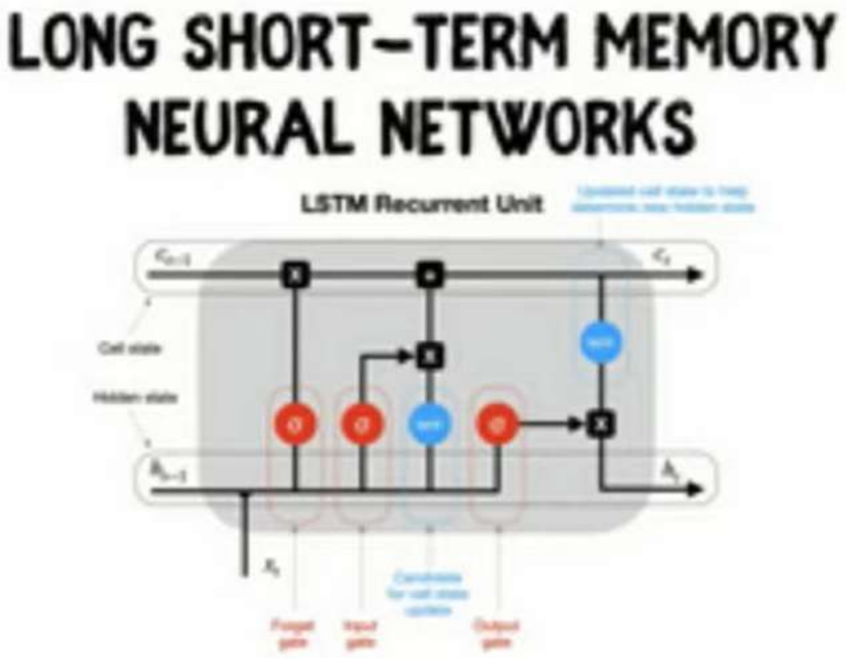
The big difference is that **Recurrent Neural Networks** also have **feedback loops.**



In our example, **The Vanishing/Exploding Gradient** problem has to do with the **weight** along the squiggle that we copy each time we **unroll** the network.



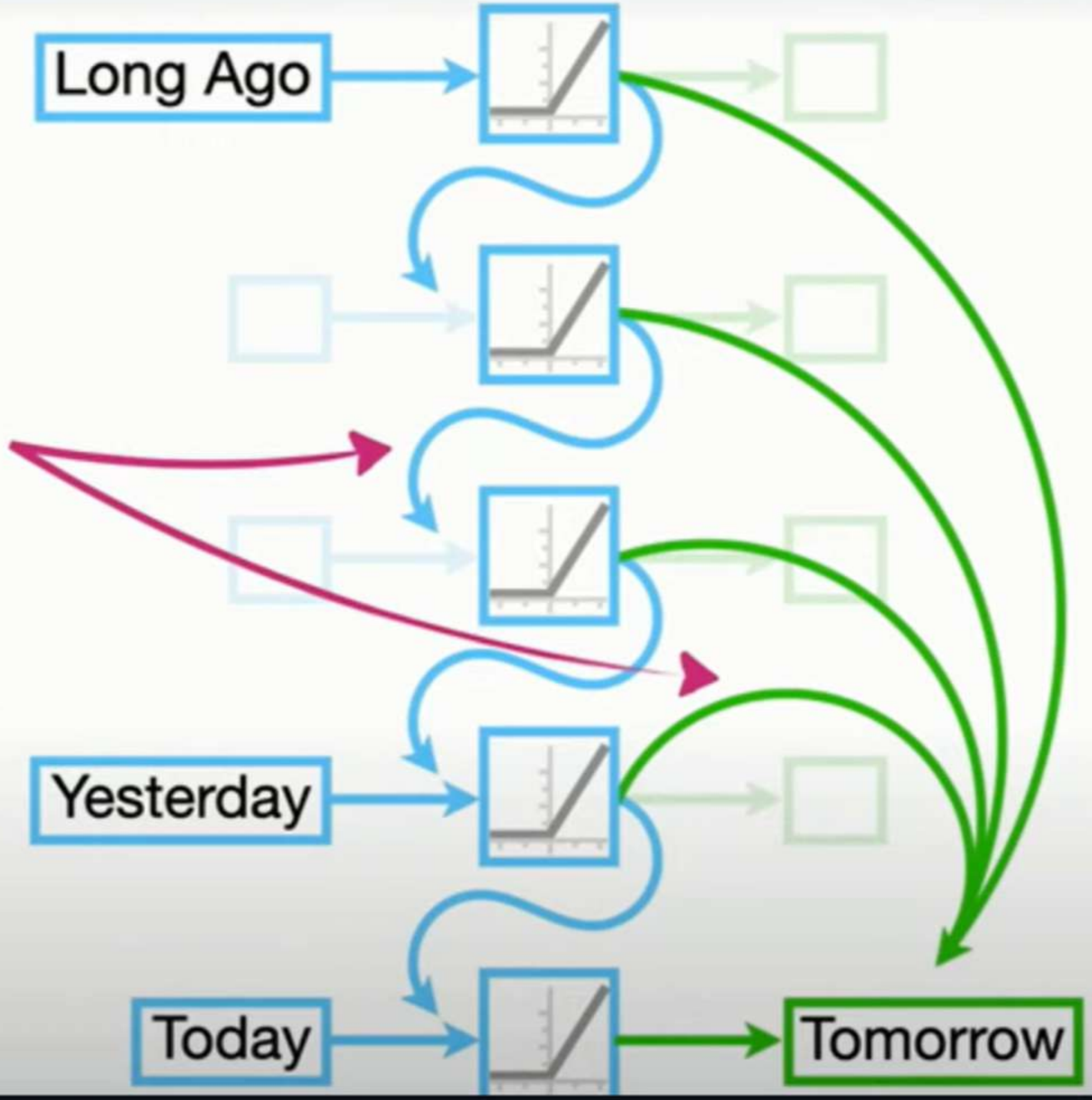
Generative AI History:
Long Short Term Memory - 1997



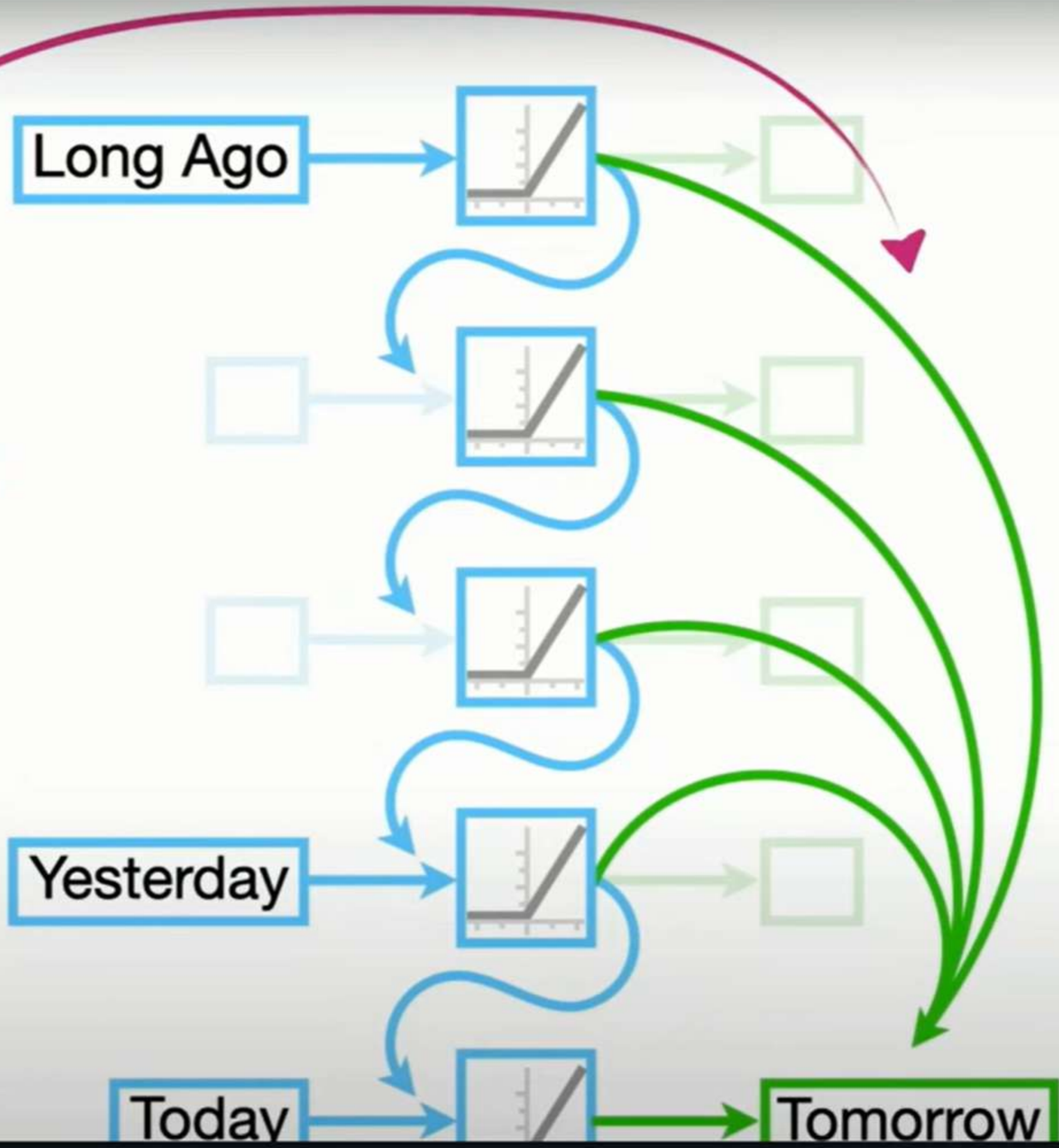
LSTMs are a specialized type of RNN designed to address the vanishing gradient problem, which occurs when training RNNs on long sequences.

→

…**Long Short-Term Memory** uses two separate paths to make predictions about tomorrow.

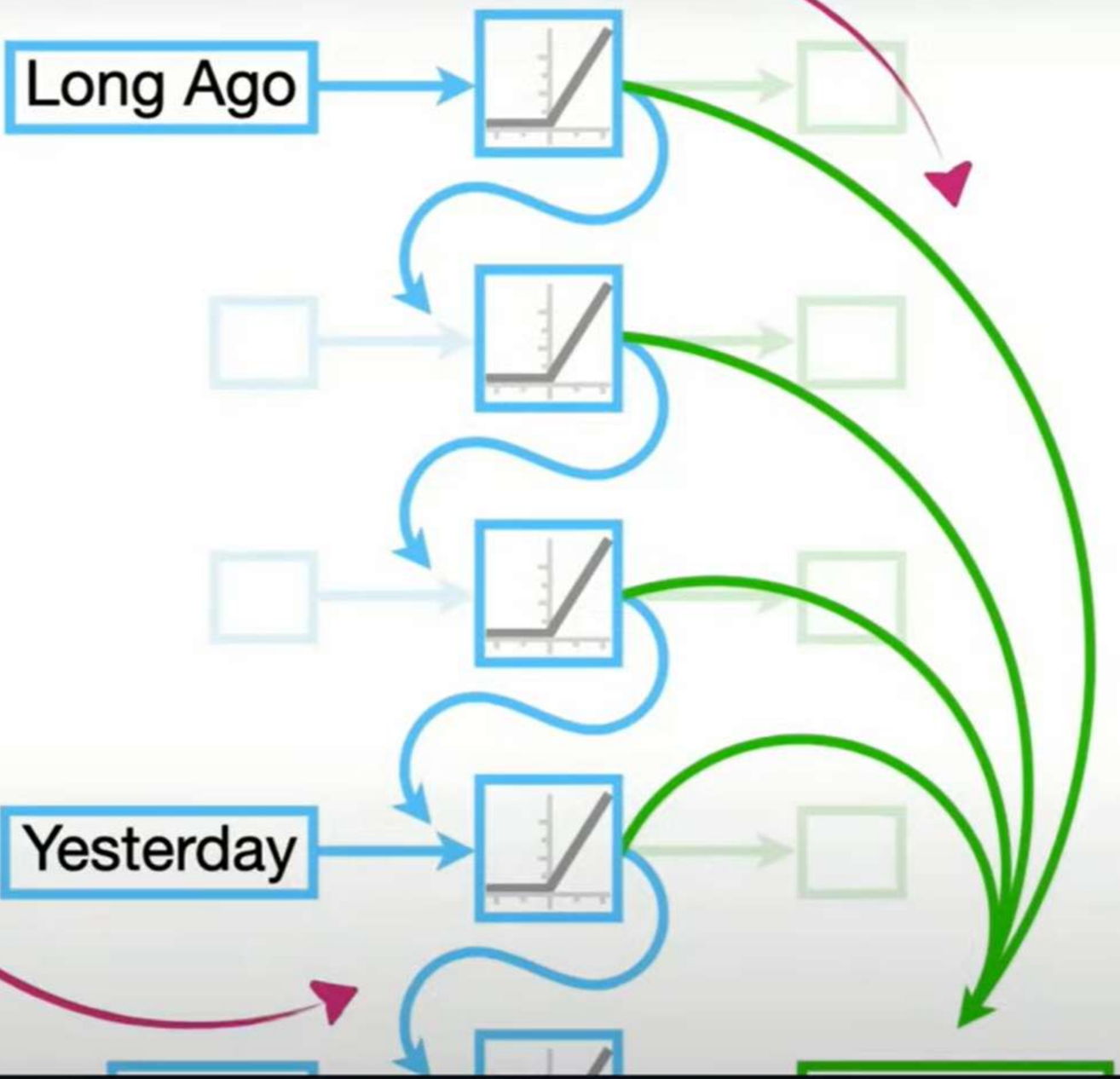


One path is for **Long-Term Memories**…

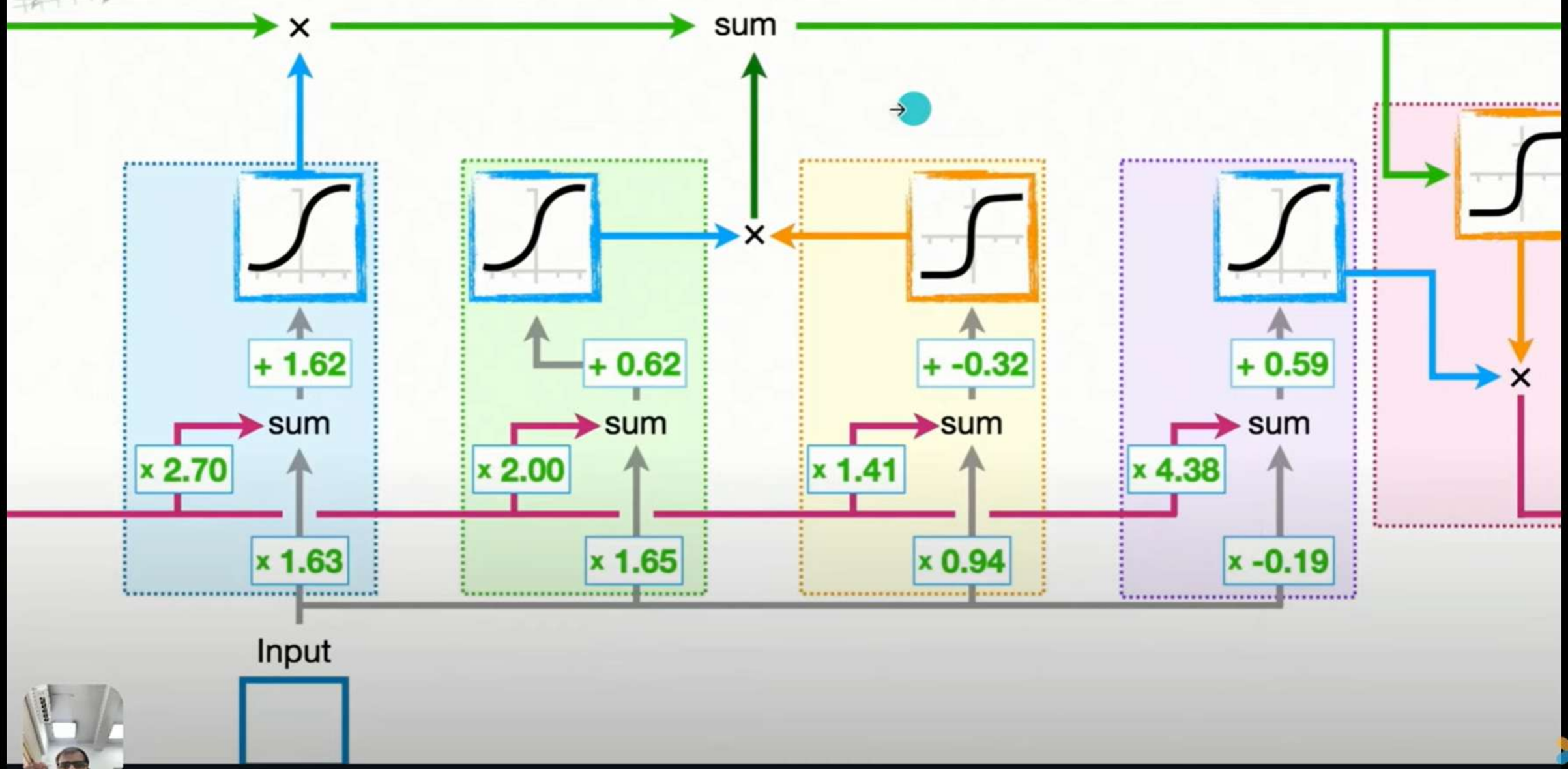


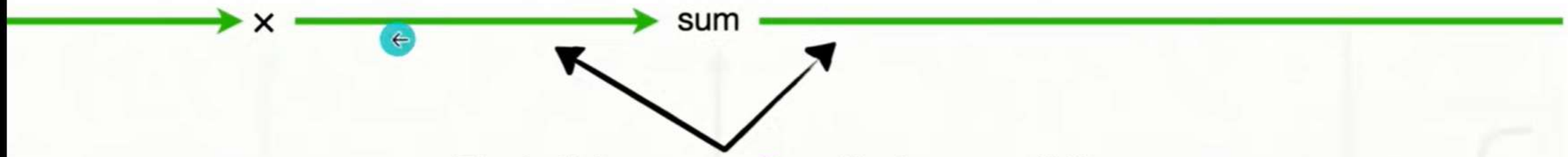
One path is for **Long-Term Memories**…

…and one is for **Short-Term Memories**.

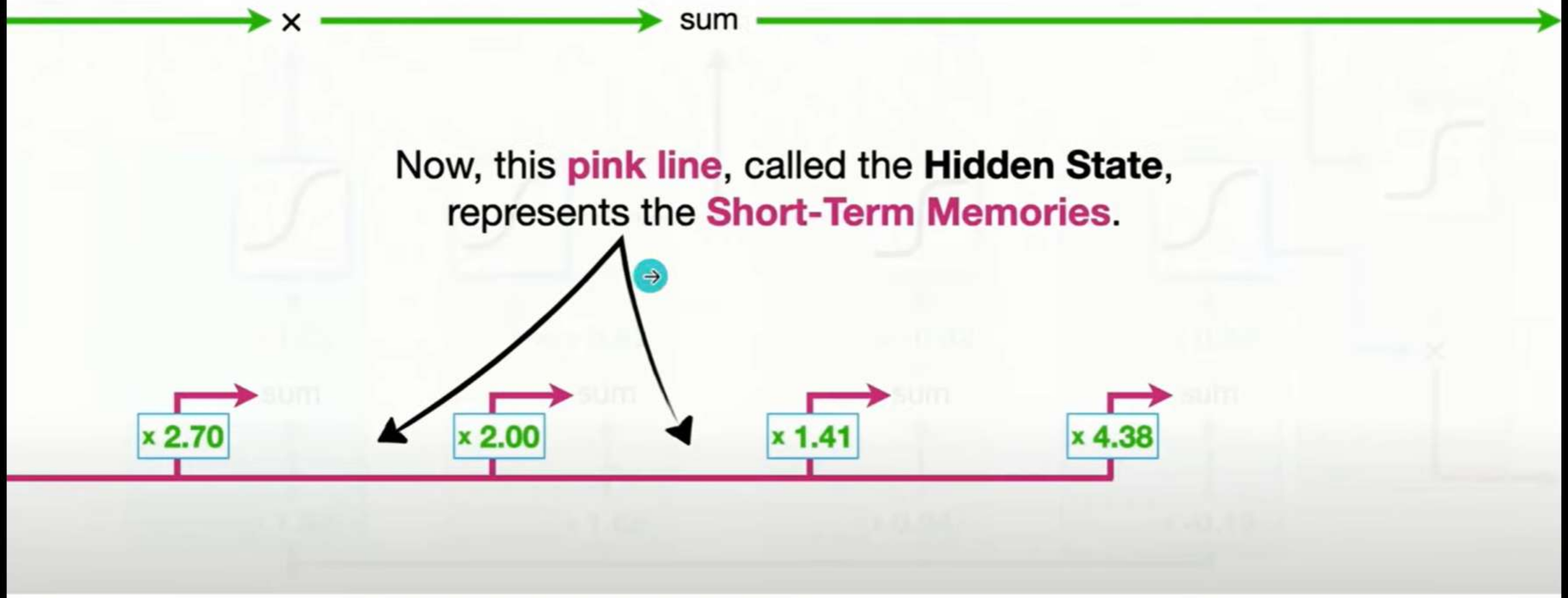


…let's talk about how the **Long Short-Term Memory** unit works.

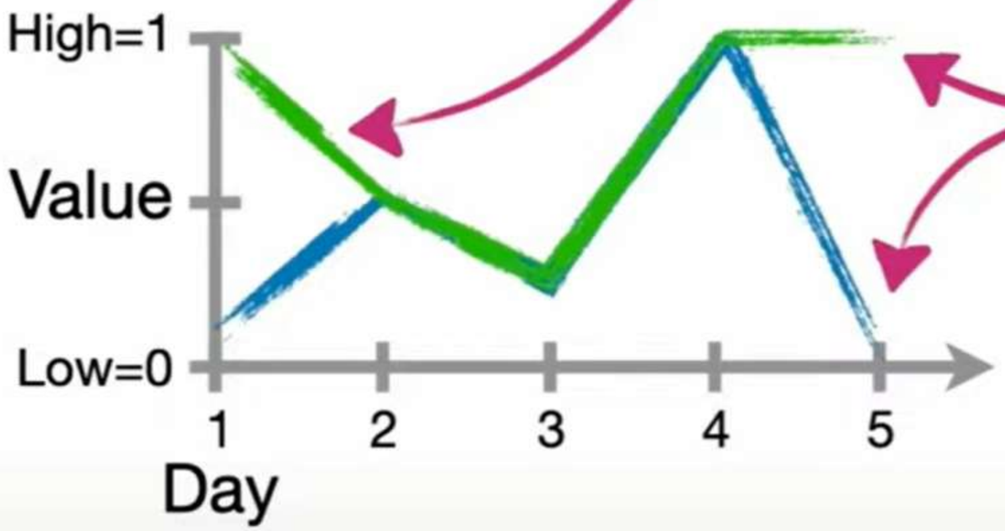




First, this **green line** that runs all the way across the top of the unit is called the **Cell State** and represents the **Long-Term Memory**.

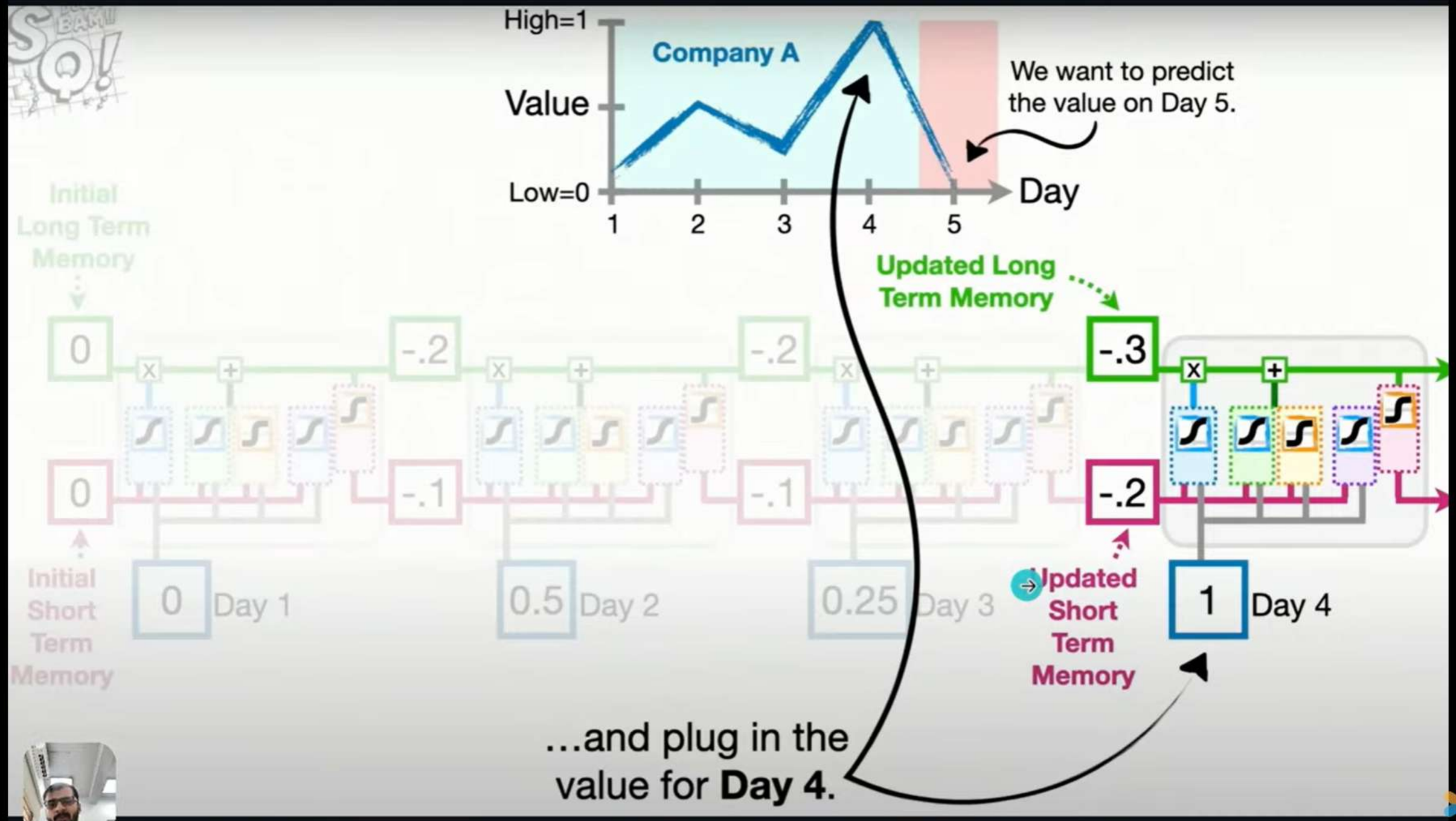

x
sum
Now, this pink line, called the Hidden State,
represents the Short-Term Memories.
x 2.70
x 2.00
x 1.41
x 4.38

Given this sequential data, we want the **LSTM** to remember what happened on **Day 1**…



…so it can correctly predict what will happen on **Day 5**.

Company A =


High=1
Value
Low=0
Company A
1
2
3
4
5
Day
We want to predict the value on Day 5.
Initial Long Term Memory
Updated Long Term Memory
0
-.2
-.2
-.3
0
-.1
-.1
-.2
Initial Short Term Memory
0
Day 1
0.5
Day 2
0.25
Day 3
Updated Short Term Memory
1
Day 4
…and plug in the value for Day 4.