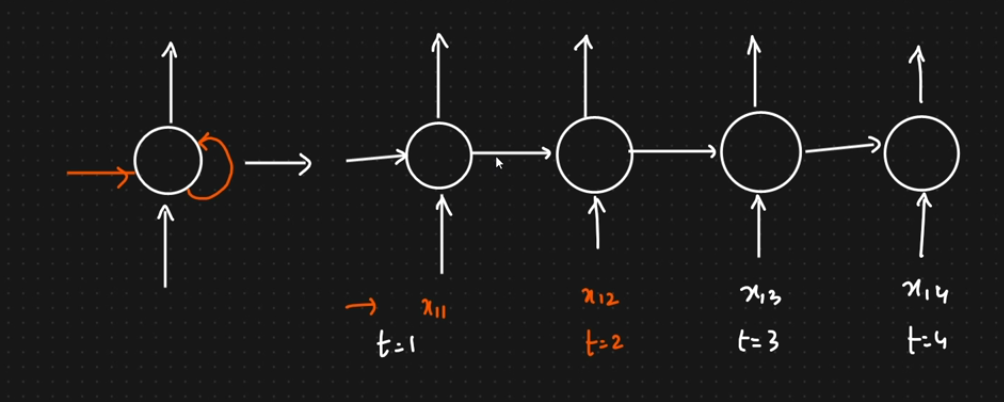
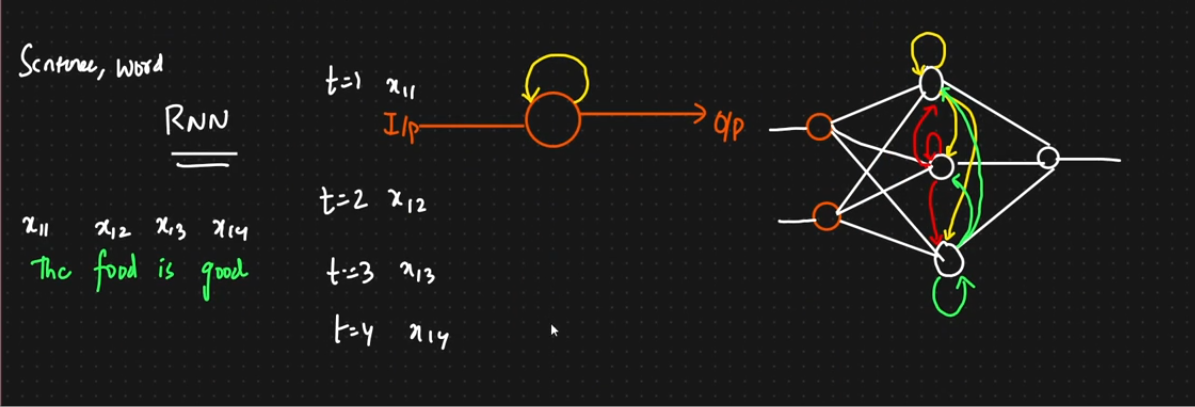
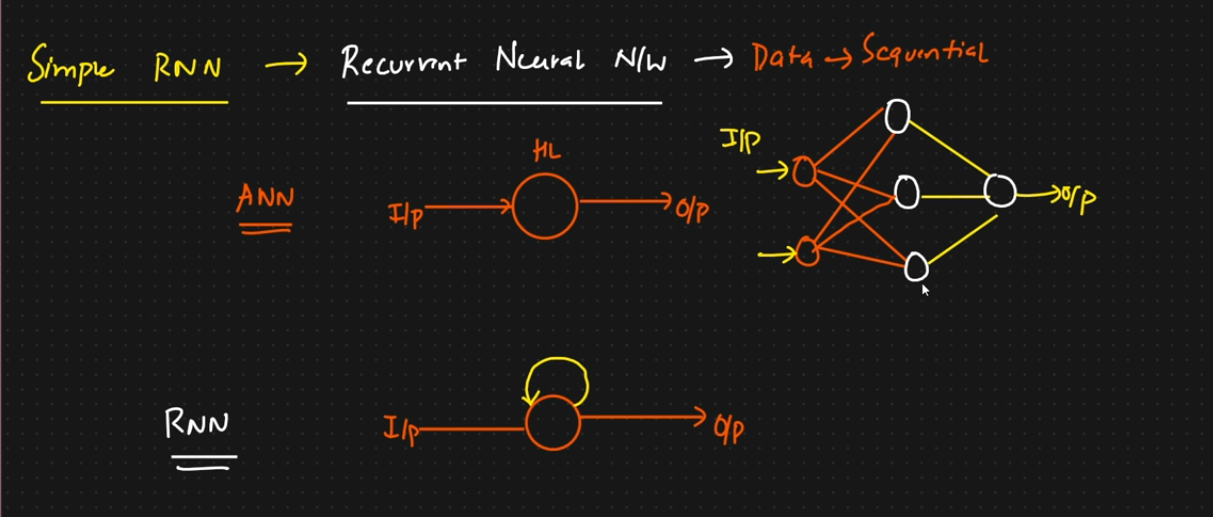
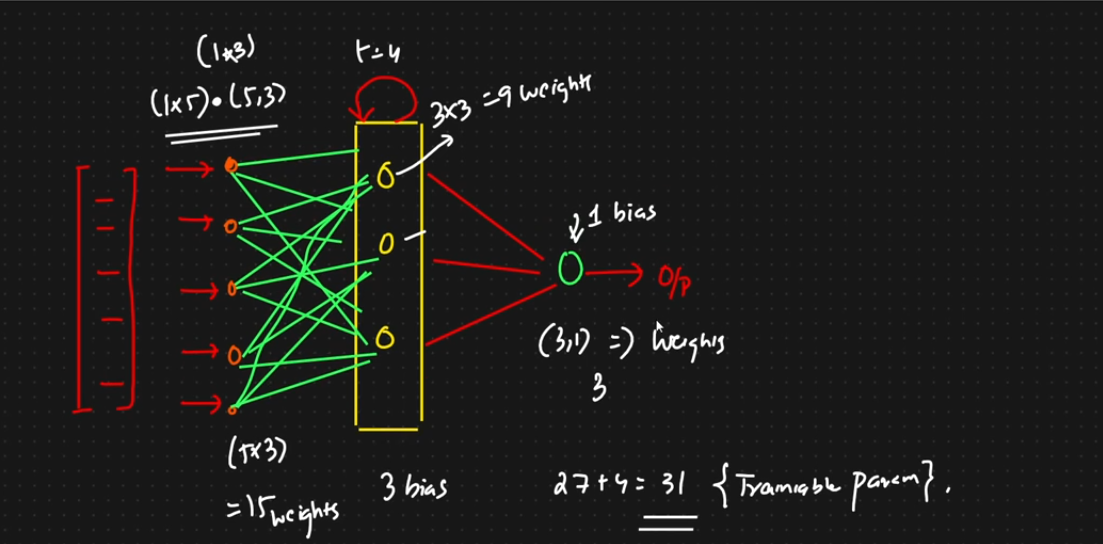
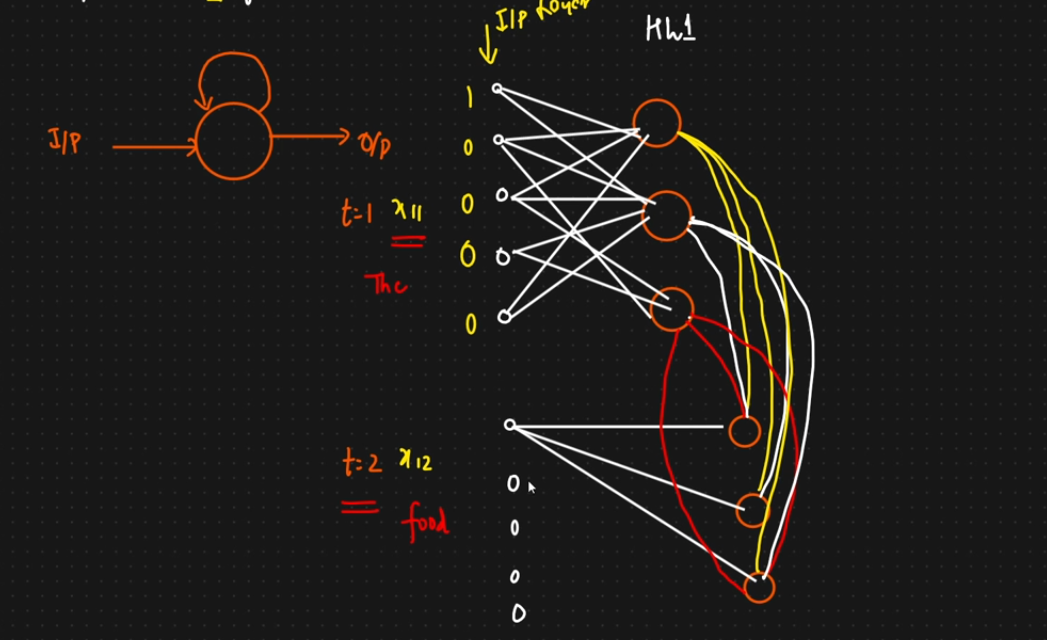
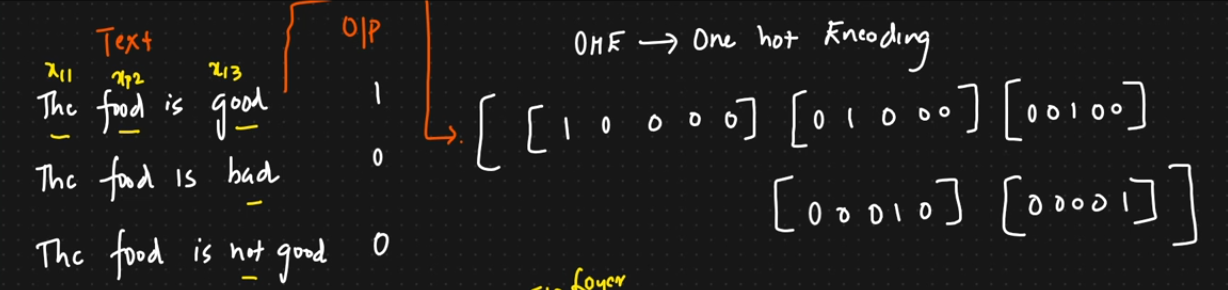
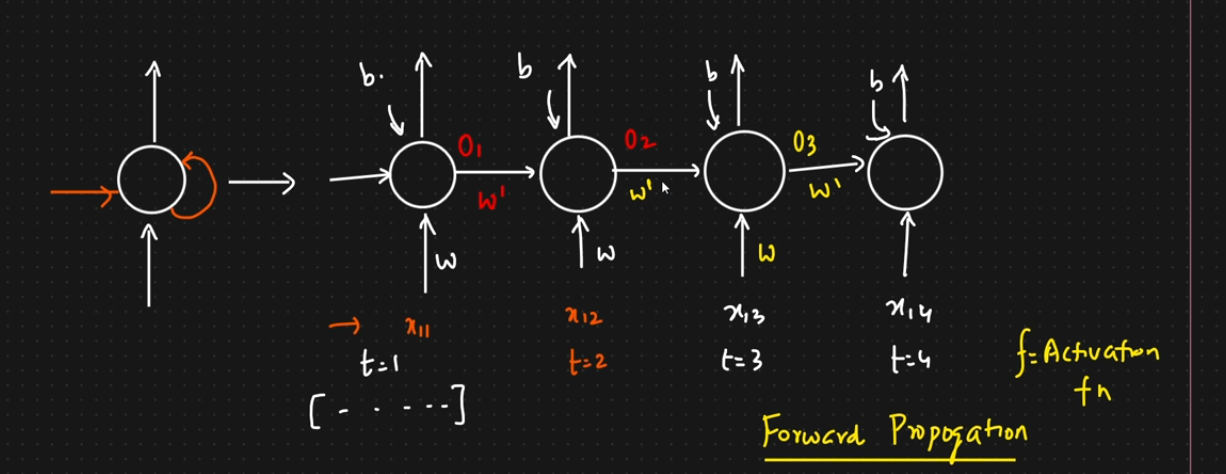
Simple ANN vs Simple RNN.

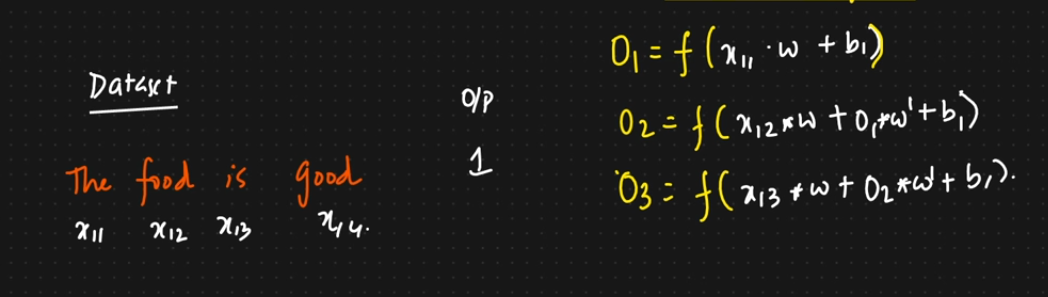


Simple RNN

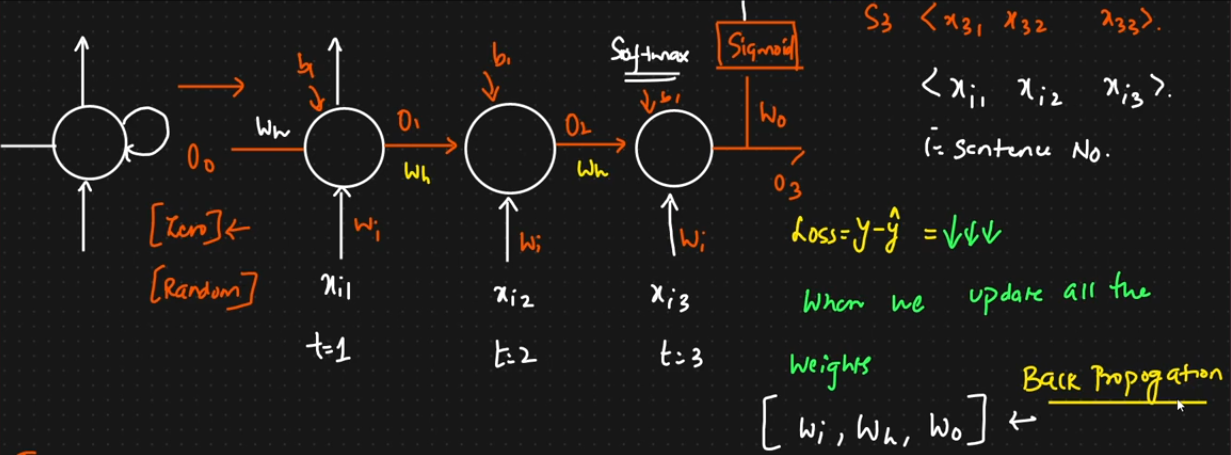


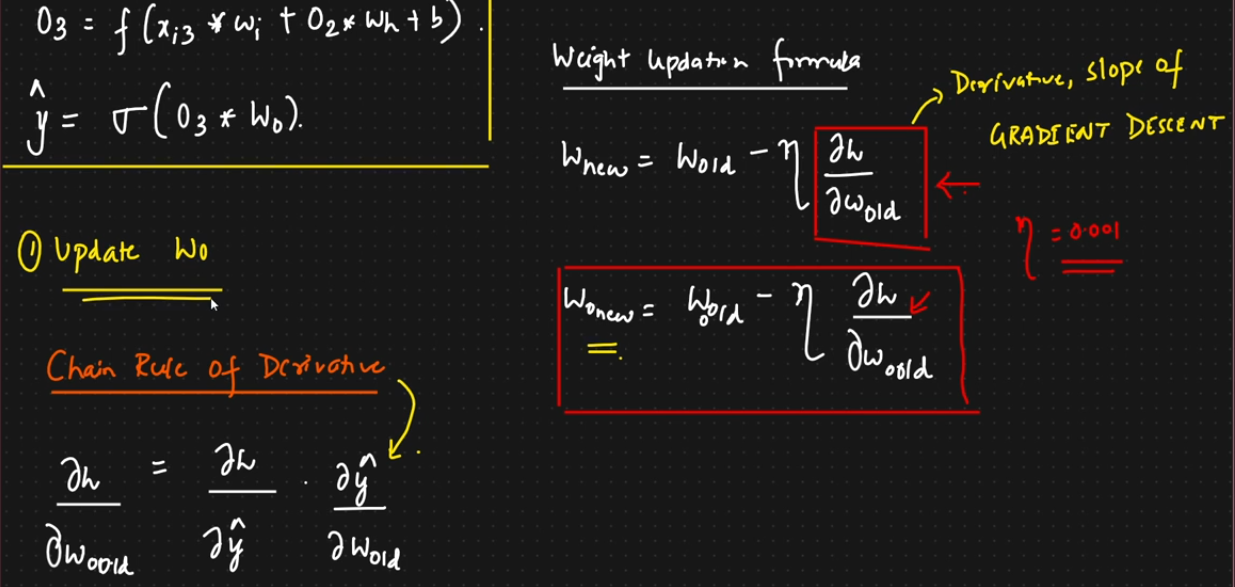
Forward Propogation.

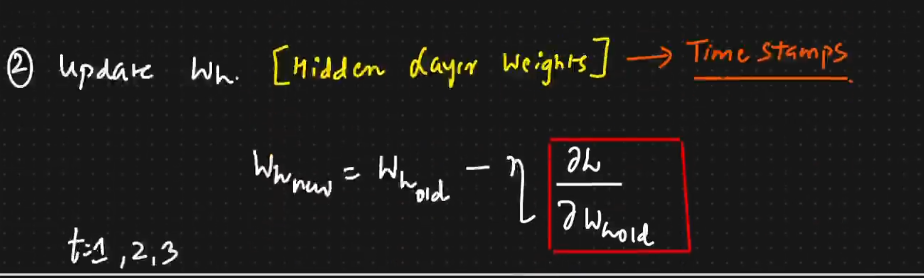


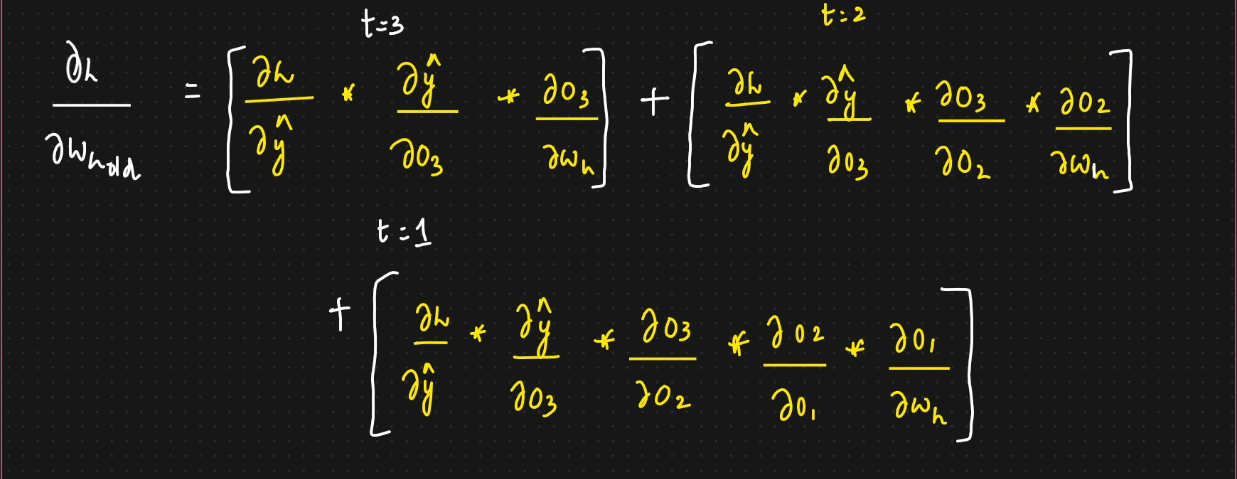


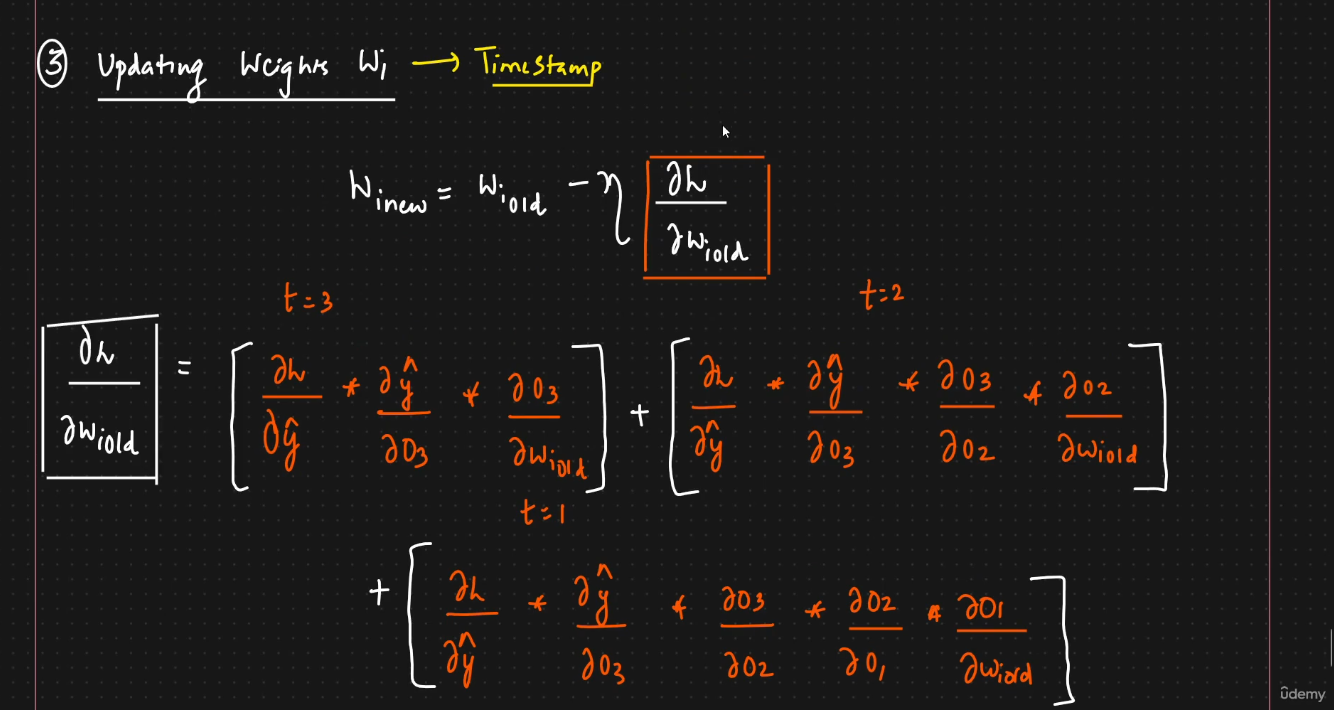
Back Propogation.











Problems with RNN:

1. Long term dependency cannot be captured.
2. Vanishing Gradient problem (Suppose len(sentence) is 100 (chain rule will be big), then the words that occurred earlier will not contribute to updating weights).

LSTM.

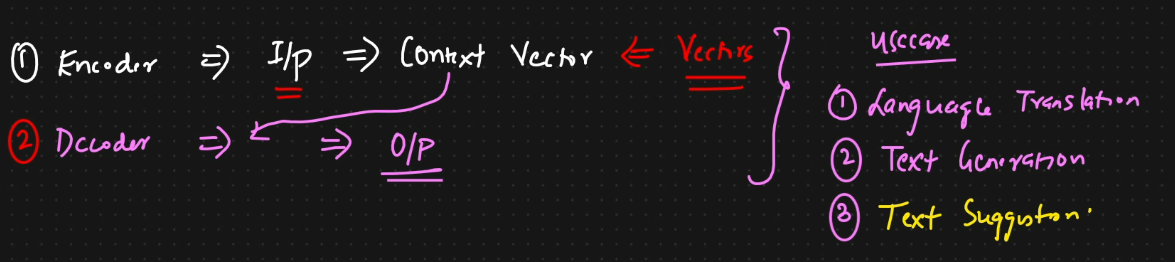
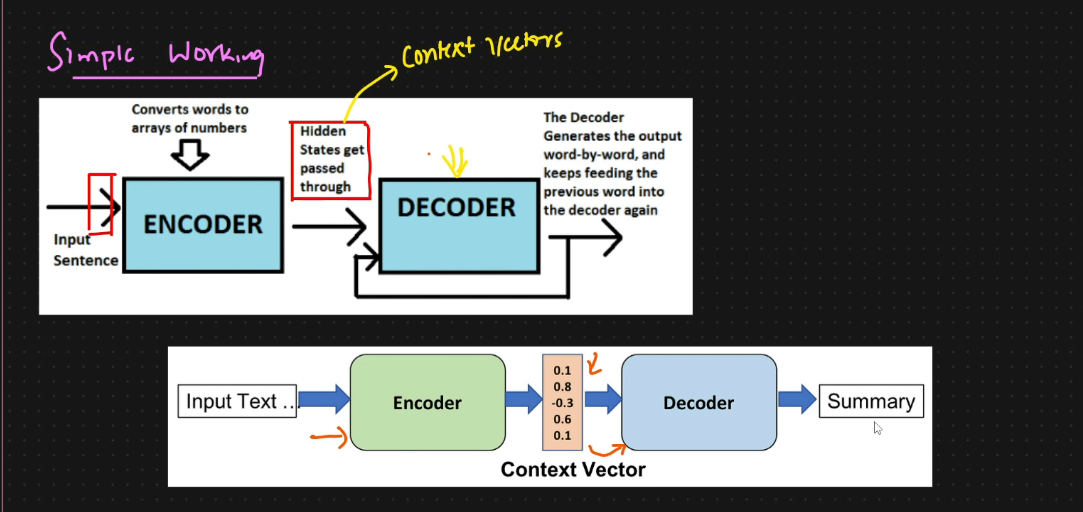
Statquest.

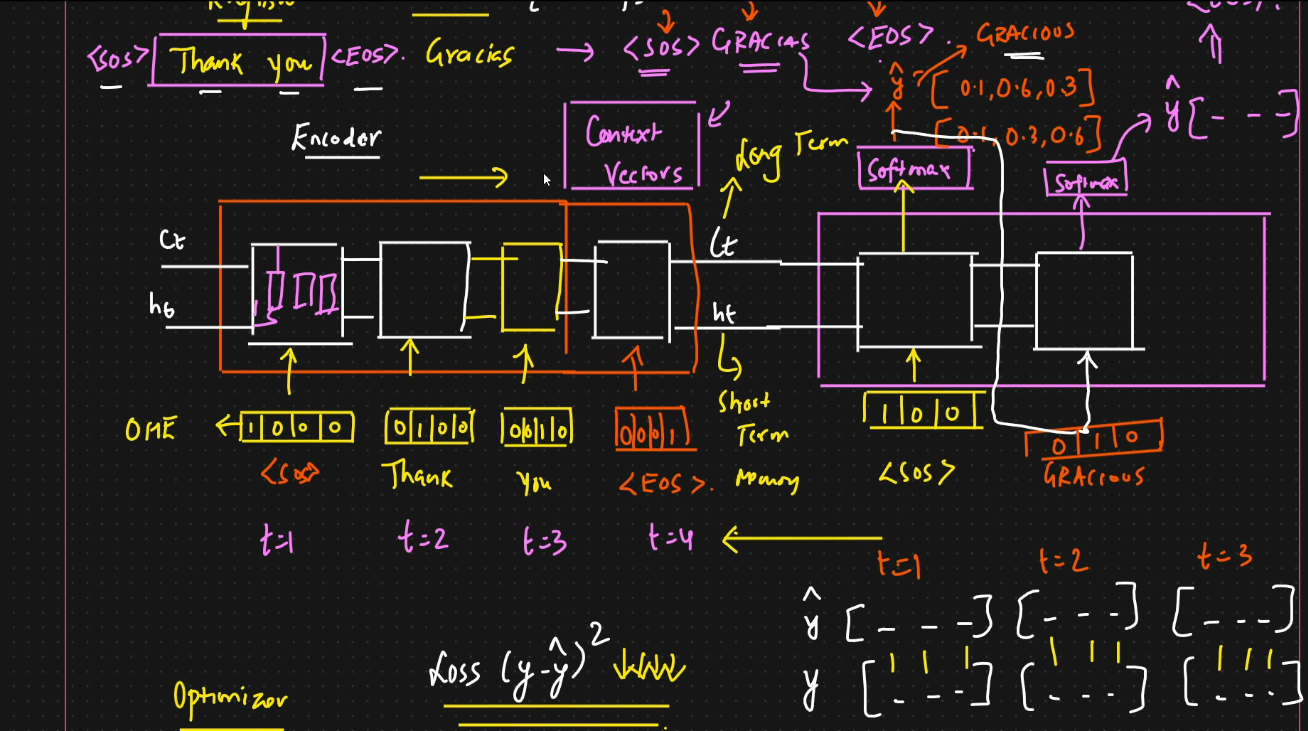
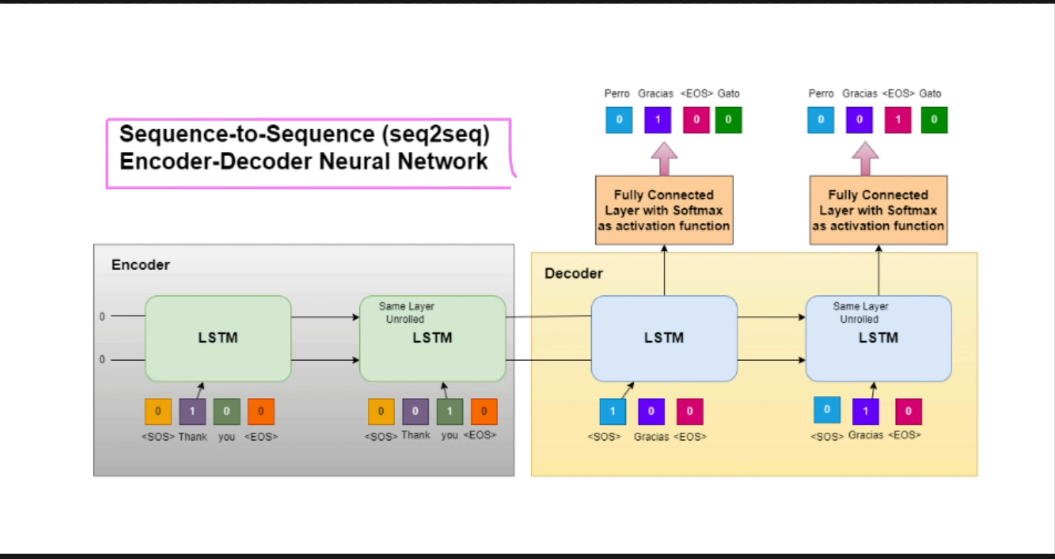
GRU.

Simplified version of LSTM with less params.

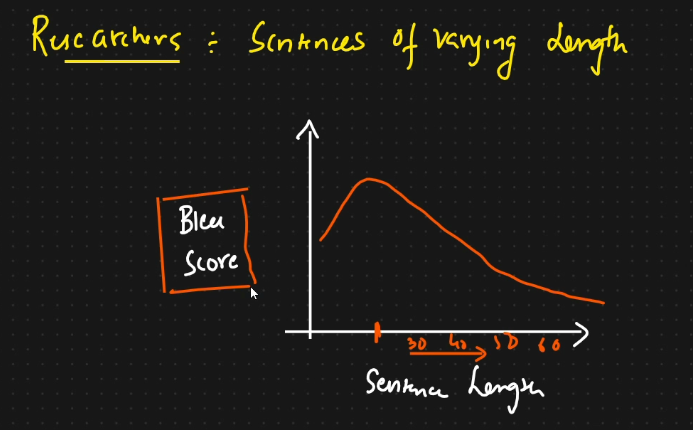
Will try to update it in future.

Encoder Decoder Seq2Seq.





Problems:



Attention in Seq2Seq:

