## horizontal line

Problem with RNN , Rise of LSTM

25.05.2025

RNN is specially for sequential data like textual , time series etc.

But RNNs suffer from 2 major problems :

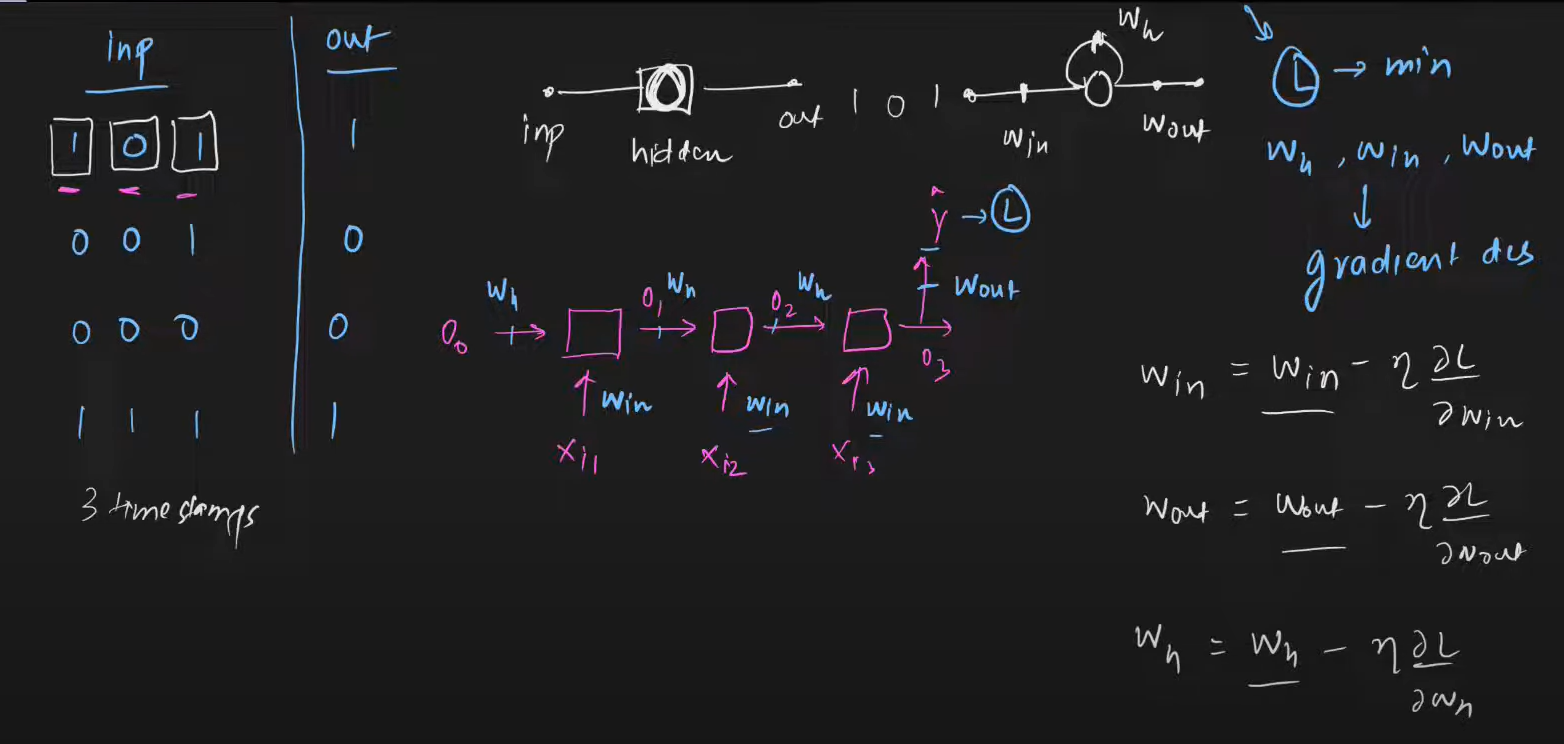
* Problem of long term dependency
* Stagnant Training

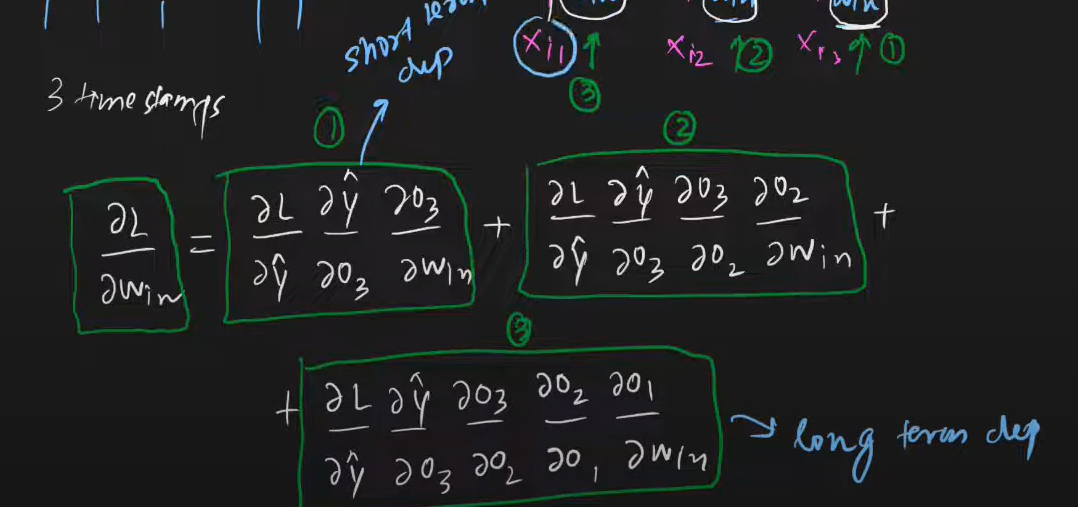
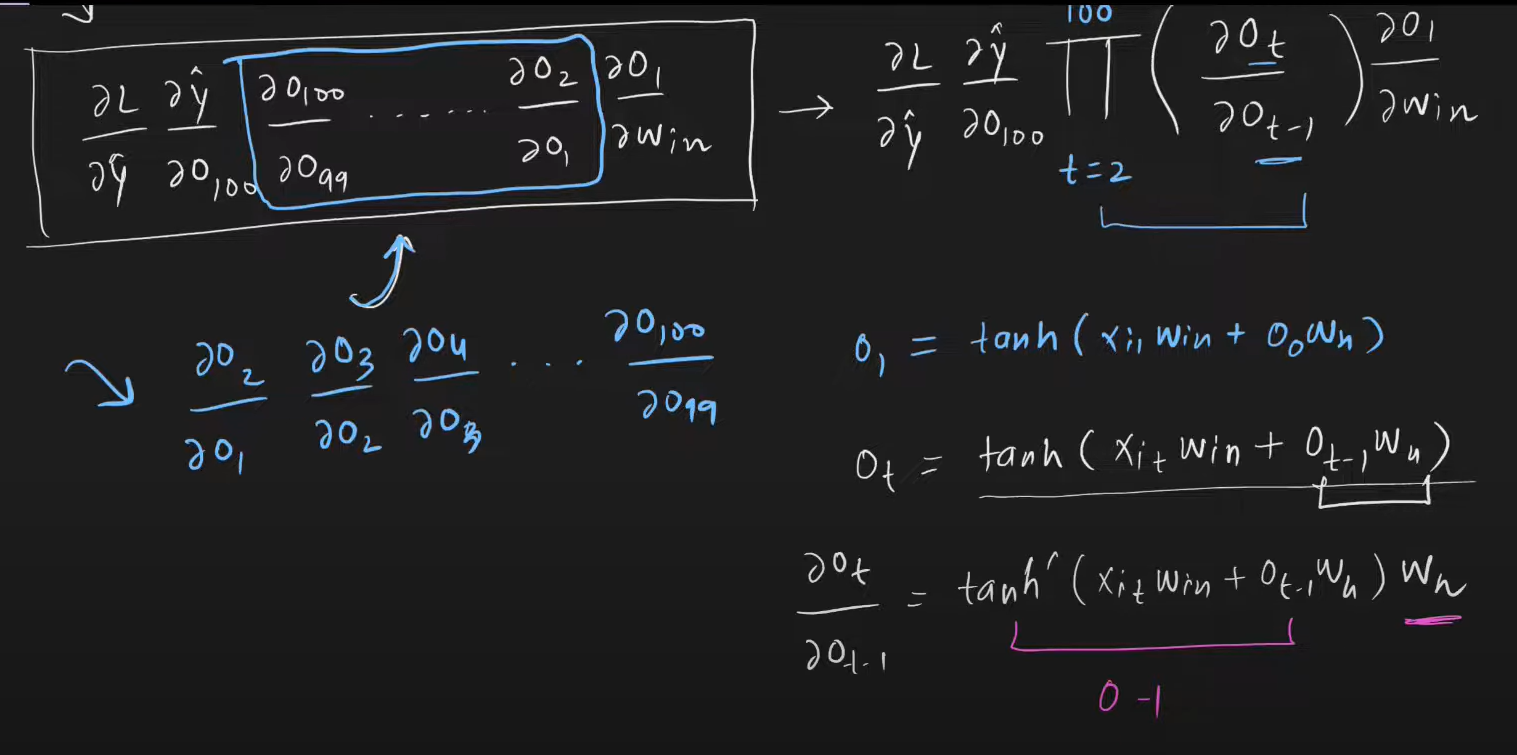
Both of these problems arise due to Unstable gradient.

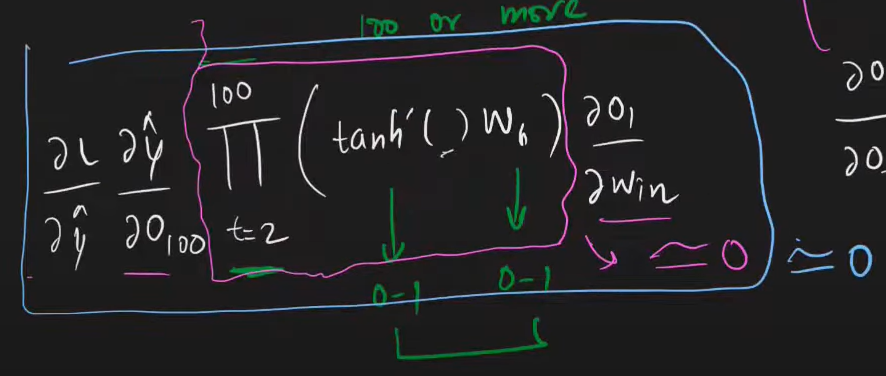
E.g. Maharashtra is \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ the language spoken here is \_\_\_\_\_\_\_\_\_\_.

To answer this one needs to know the first word which is comparatively a long term dependency where RNNs perform poorer while good in Short terms.

## #Problem 1 : Problem of long term dependency







More than 100 terms of it would give a sum approaches to zero

That leads to “Vanishing Gradient Problem”.

Solutions to this :

* Use different activation function like relu / leaky relu.
* Better weight Initialization
* Skip rnns
* LSTM

## #Problem2 : Stagnant Training (Exploding Gradients)

Same reverse as of problem1

When using relu as activation function the sum of all those summarize to a large sum leads to unstable gradient and training stops.

Solutions to this :

* Gradient Clipping (stop gradients training at specific value)
* Controlled learning rate
* LSTM