Problem with RNN

19 December 2022 16:33

RNNs → sequential data → textual, time series

suffer (2) major problems → LTSMs (Why)

RNN

→ Problem of long term dependency
→ ~~Unstable~~ gradients Stagnated values → Unstable gradients

Applicat next word predicts

→ RNN

timesteps

Marathi is spoken in Maha

short

RNN

Vanishing gradient prob → long term → RNN → keypad

Maha is a beautiful place ± went there last year, but I could not enjoyed proper bcz ± dont understa marathi

Exploding gradient problem → RNN → possible sol(n) → remove

Problem #1 → Problem of long term dependency → Vanishing

inp	out
1 0 1	1
0 0 1	0
0 0 0	0
1 1 1	1

3 time stamps

short term dep

0-1

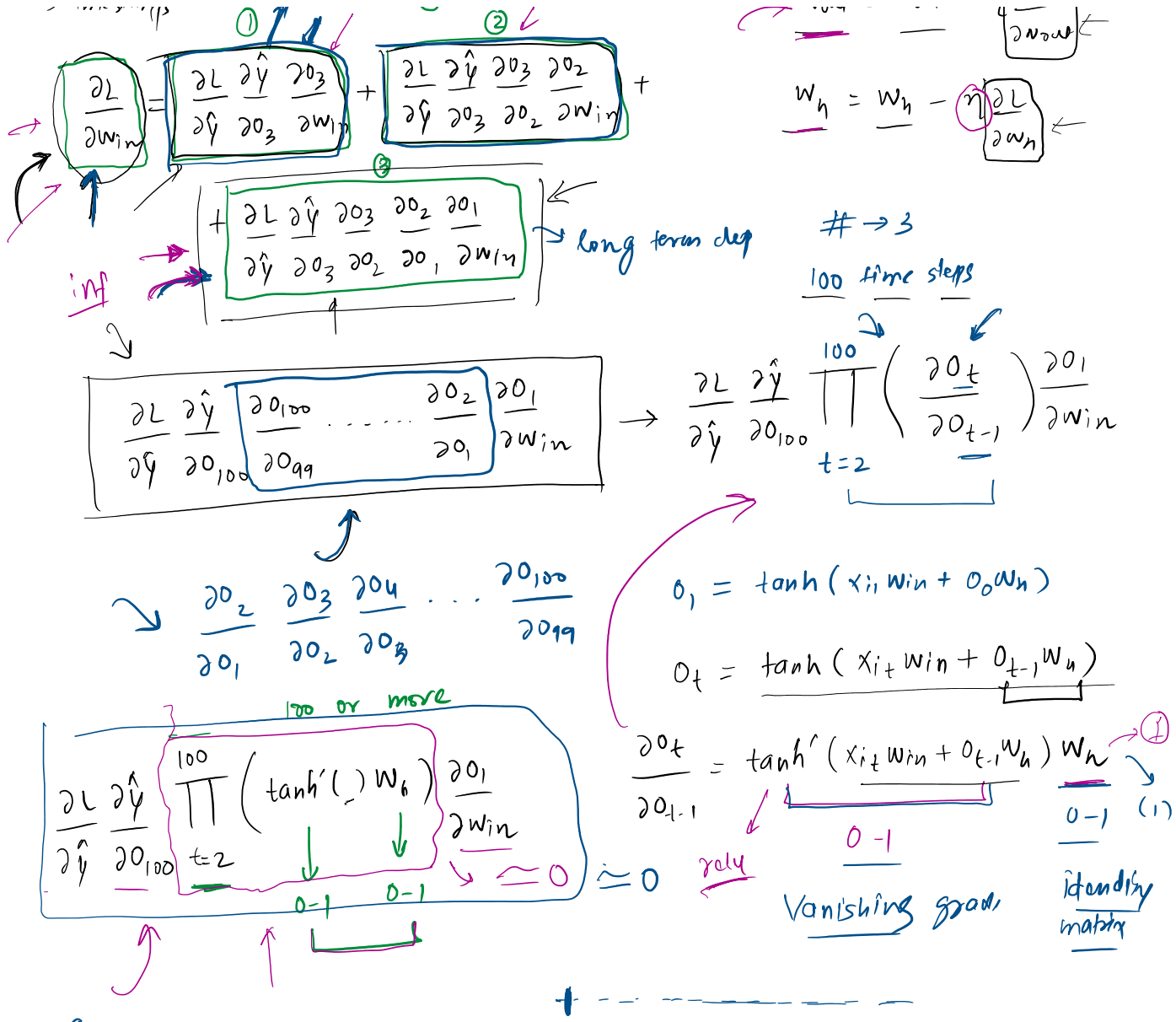
inp hidden out 1 0 1 W_{in} W_h W_{out}

L → min

W_h, W_{in}, W_{out} → gradient des

$$W_{in} = W_{in} - \eta \frac{\partial L}{\partial W_{in}}$$

$$W_{out} = W_{out} - \eta \frac{\partial L}{\partial W_{out}}$$


Long term dep
#→3
100 time steps
100 or more
Vanishing grad.
identity matrix
relu
Soln
1) Diff activation → relu / leaky relu
2) Better weight init
3) Skip conns
4) LSTM
Problem #2 → Unstable Training (Exploding gradients)
1) Gradient Clipping
2) Controlled learning rate
3) LSTM

Recap

21 August 2023 11:55