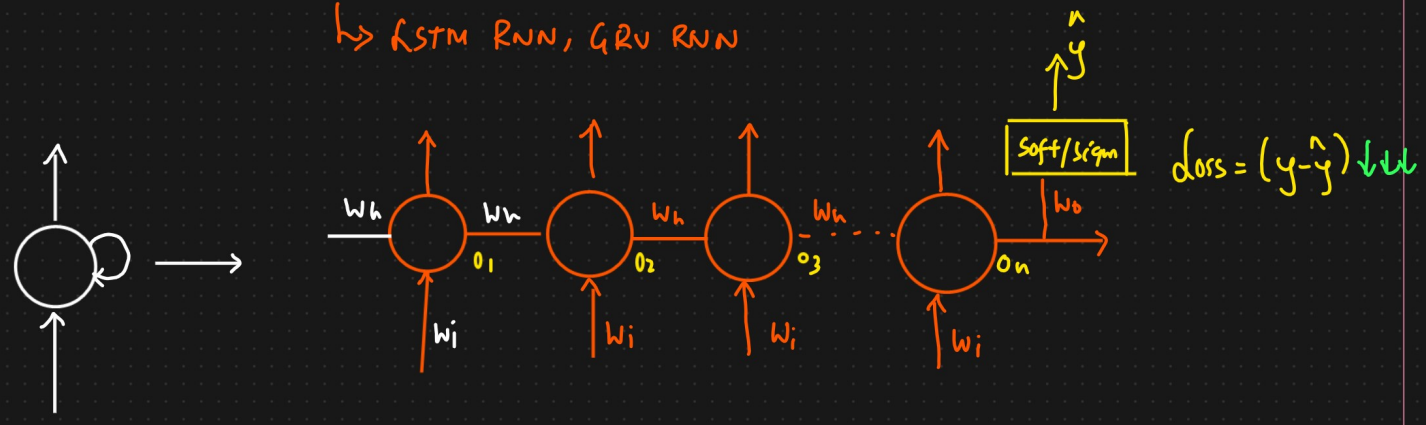
Problems With RNN

↳ LSTM RNN, GRU RNN

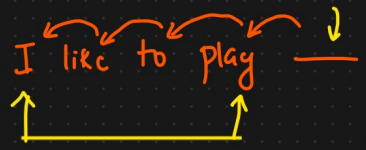


ANN → Vanishing Gradient problem.

Text	O/P
The food is good	1
The food is bad	0

Sentences → small

Text Generation



{ Sentence is . long

t:1 t=2 t=3 t=4

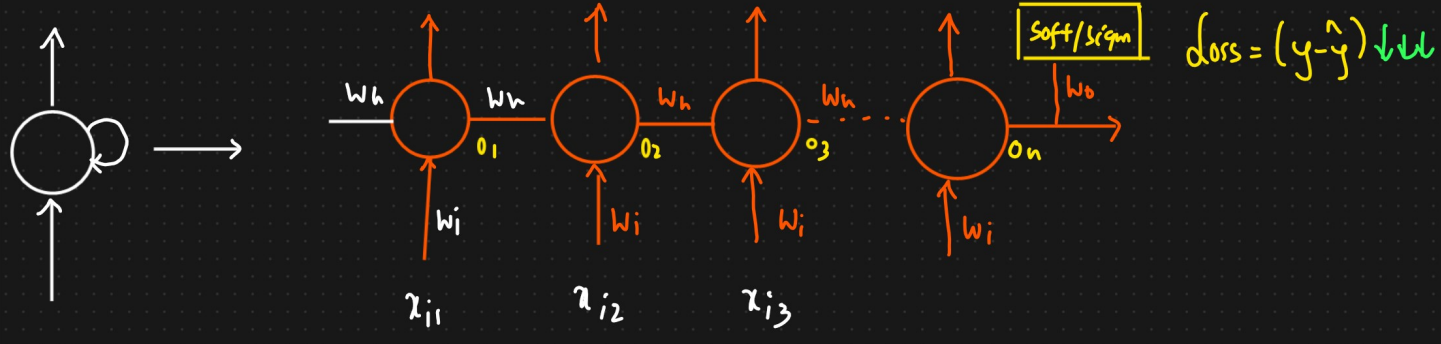
My Name is KRISH and I like sports like CRICKET, VOLLEYBALL AND ALSO LIKE TO MAKE

t=16

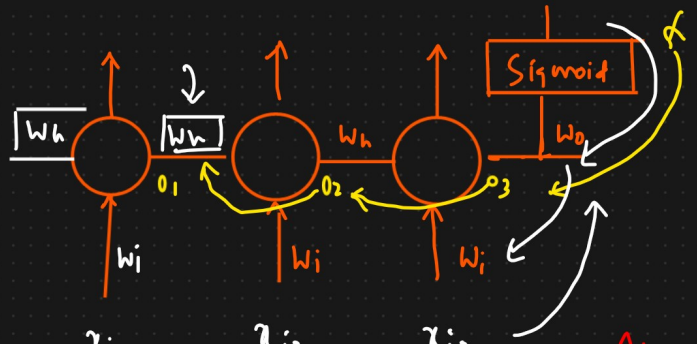
food. Videos

100 words

① long term dependency cannot be captured by RNN → Provide Accuracy ⇓ Dependency.



ŷ loss = (y-ŷ)



Updates W_h, W_i, W_o

$$W_{new} = W_{old} - \eta \frac{\partial L}{\partial W_{old}}$$

x_{i1} x_{i2} x_{i3}

$t=1$ $t=2$ $t=3$

$$\frac{\partial L}{\partial W_o} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial W_o} \Leftarrow \quad t=2$$

$t=1$

$$\frac{\partial L}{\partial W_{h_{old}}} = \left[\frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial O_3} \cdot \frac{\partial O_3}{\partial O_2} \cdot \frac{\partial O_2}{\partial W_{h_{old}}} \right] + \cdots$$

Length of sentence $\Rightarrow$ 50 words

$t=1$ $t=2$

$$\frac{\partial L}{\partial W_{h_{old}}} = \left[\frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial O_{50}} \cdot \frac{\partial O_{50}}{\partial O_{49}} \cdot \frac{\partial O_{49}}{\partial O_{48}} \cdots \frac{\partial O_2}{\partial W_{h_{old}}} \right] + \cdots$$

$0.25 \times 0.02 \times 0.01 \times 0.015$

$O_3 = \sigma(x_{i3} * W_i + O_2 * W_h + b)$

$\Downarrow$

Small Value ≈ 0

$$\frac{\partial O_3}{\partial O_2} = \frac{\partial \sigma(x_{i3} * W_i + O_2 * W_h + b) \Leftarrow O_3}{\partial O_2}$$

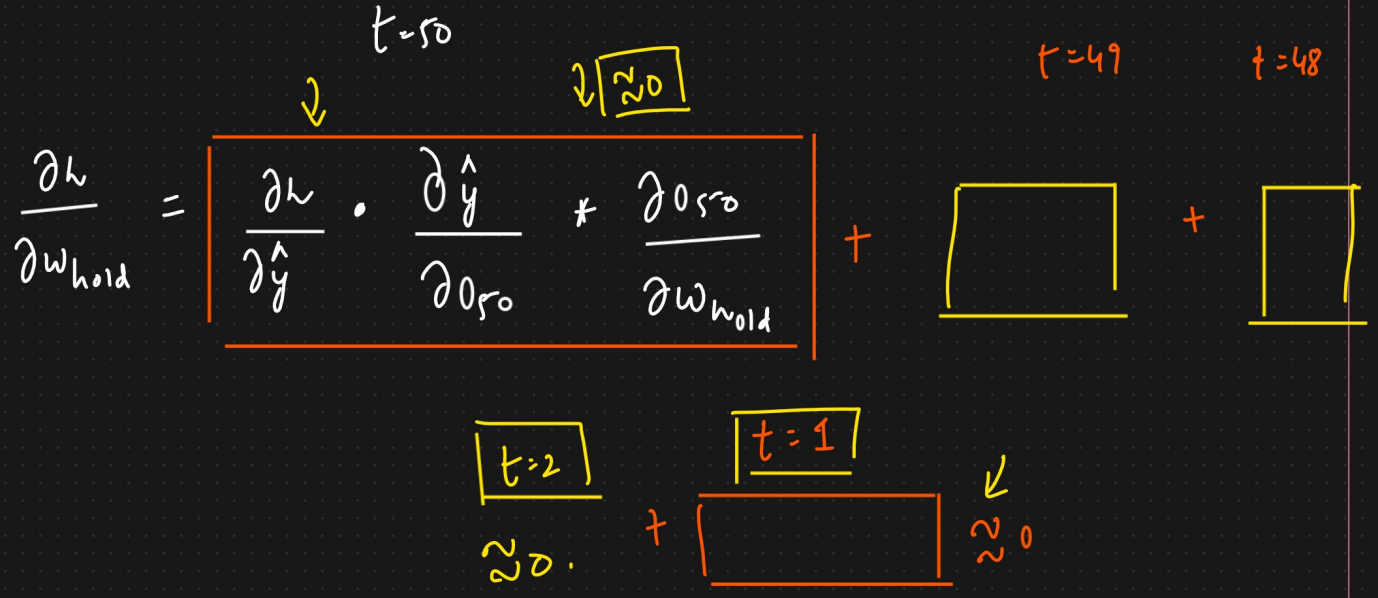
Derivative of sigmoid $0 - 0.25$

$$= \sigma'(1 * W_h) \Rightarrow [0 - 0.25] * (W_h)$$

Vanishing Gradient problem

$t=1 \Rightarrow$ The word is not participating to update the weights value.

Chain Rule is Big



① Relu, Leaky Relu →

② LSTM RNN → Long Short term Memory RNN } ⇒ Simple RNN.

③ GRU RNN →