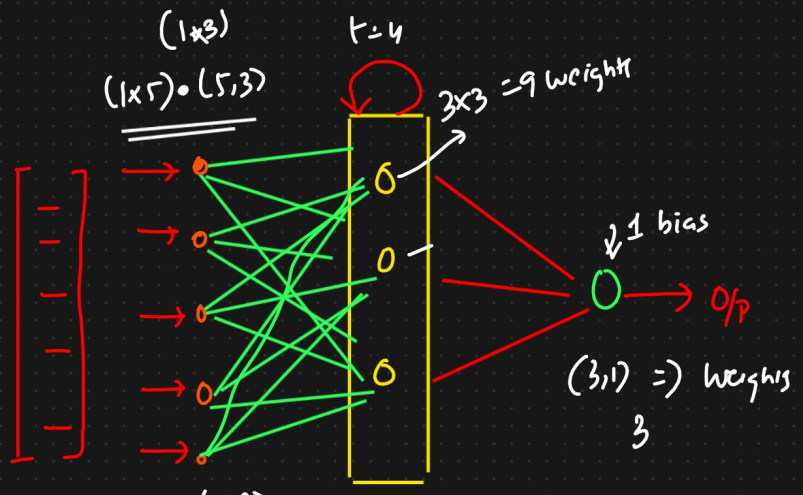
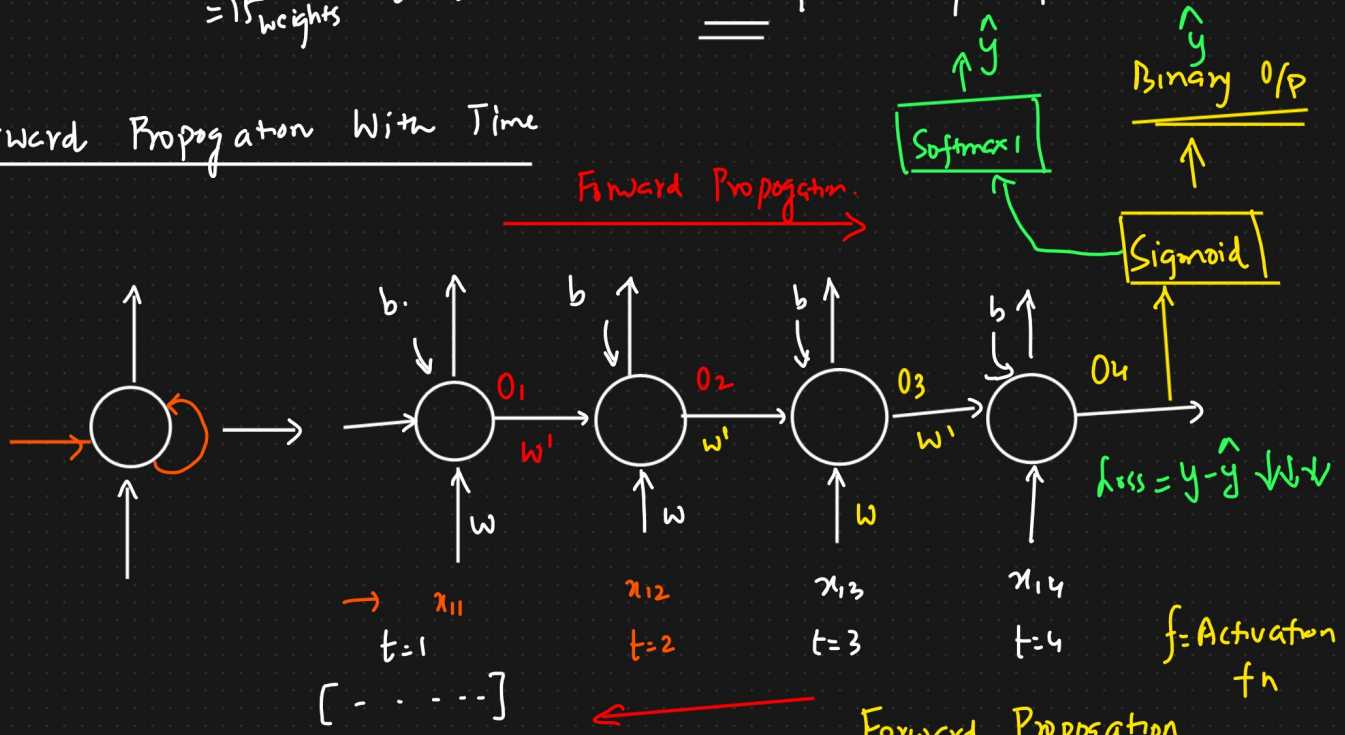


$27 + 4 = 31$ {Trainable params}.

Forward Propogation With Time



Forward Propogation

$O_1 = f(x_{11} \cdot w + b_1)$

$O_2 = f(x_{12} * w + O_1 * w' + b_1)$

$O_3 = f(x_{13} * w + O_2 * w' + b_1)$.

⋮

O_n

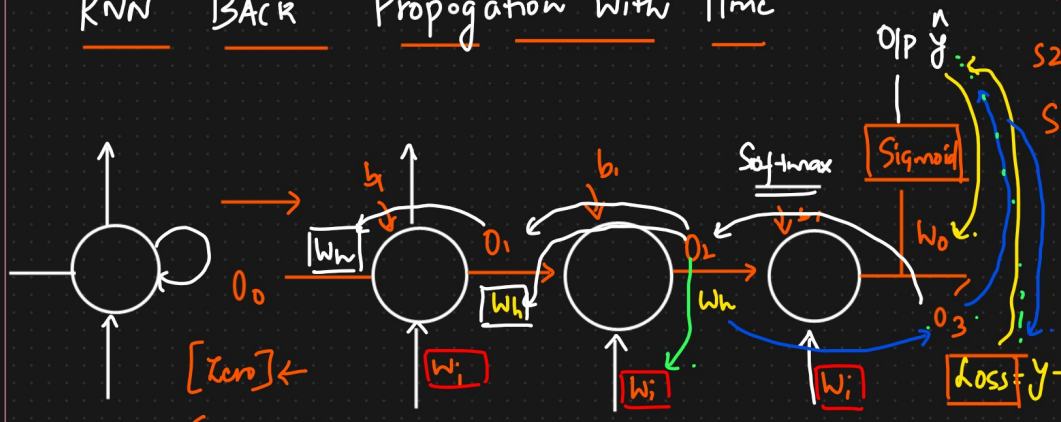
Dataset

The	food	is	good
x_{11}	x_{12}	x_{13}	x_{14}.

O/p

1

RNN BACK Propogation With Time



O/P

S1 ⟨x11 x12 x13⟩ 1

S2 ⟨x21 x22 x23⟩ 0

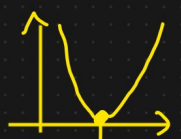
S3 ⟨x31 x32 x33⟩.

⟨xi1 xi2 xi3⟩.

i = Sentence No.

When we update all the weights

[Wi , Wh , Wo] ← Back Propogation



Forward Propogation

$$O_1 = f(x_{i1} * W_i + O_0 W_h + b_1)$$

$$O_2 = f(x_{i2} * W_i + O_1 * W_h + b)$$

$$O_3 = f(x_{i3} * W_i + O_2 * W_h + b).$$

$$\hat{y} = \sigma(O_3 * W_o).$$

Backward Propagation with Time

Update [Wi, Wh, Wo]

Weight updation formula

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

Derivative, slope of GRADIENT DESCENT

$\eta = 0.001$

$$W_{onew} = W_{oold} - \eta \frac{\partial h}{\partial w_{oold}}$$

① Update Wo

Chain Rule of Derivative

$$\frac{\partial h}{\partial w_{oold}} = \frac{\partial h}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial w_{old}}$$

② Update Wh. [Hidden layer weights] → Time Stamps

$$W_{h_{new}} = W_{h_{old}} - \eta \frac{\partial h}{\partial W_{hold}}$$

t = 1, 2, 3

$$\Rightarrow \frac{\partial L}{\partial w_{h\,old}} = \underbrace{\left[\frac{\partial L}{\partial \hat{y}} * \frac{\partial \hat{y}}{\partial o_3} * \frac{\partial o_3}{\partial w_h}\right]}_{t=3} + \underbrace{\left[\frac{\partial L}{\partial \hat{y}} * \frac{\partial \hat{y}}{\partial o_3} * \frac{\partial o_3}{\partial o_2} * \frac{\partial o_2}{\partial w_h}\right]}_{t=2}$$

$$+ \underbrace{\left[\frac{\partial L}{\partial \hat{y}} * \frac{\partial \hat{y}}{\partial o_3} * \frac{\partial o_3}{\partial o_2} * \frac{\partial o_2}{\partial o_1} * \frac{\partial o_1}{\partial w_h}\right]}_{t=1}$$

③ Updating Weights w_i → Timestamp

$$W_{i\,new} = W_{i\,old} - \eta \boxed{\frac{\partial L}{\partial w_{i\,old}}} \quad \{\text{update } w_i \text{ weights}\}$$

$$\boxed{\frac{\partial L}{\partial w_{i\,old}}} = \underbrace{\left[\frac{\partial L}{\partial \hat{y}} * \frac{\partial \hat{y}}{\partial o_3} * \frac{\partial o_3}{\partial w_{i\,old}}\right]}_{t=3} + \underbrace{\left[\frac{\partial L}{\partial \hat{y}} * \frac{\partial \hat{y}}{\partial o_3} * \frac{\partial o_3}{\partial o_2} * \frac{\partial o_2}{\partial w_{i\,old}}\right]}_{t=2}$$

$$+ \underbrace{\left[\frac{\partial L}{\partial \hat{y}} * \frac{\partial \hat{y}}{\partial o_3} * \frac{\partial o_3}{\partial o_2} * \frac{\partial o_2}{\partial o_1} * \frac{\partial o_1}{\partial w_{i\,old}}\right]}_{t=1}$$