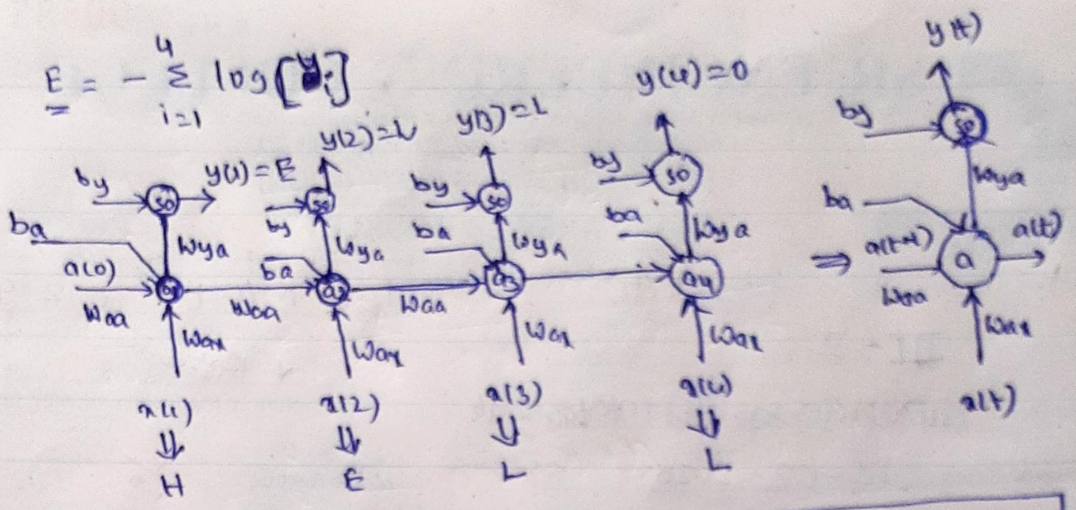
$$E = -\sum_{i=1}^{4} \log [y_i]$$



(1)

$$\Delta W_{ya} = -\eta \frac{\partial E}{\partial W_{ya}}$$

$$net_y^{(t)} = [W_{ya} * a(t) + b_y]$$

$$y(t) = \frac{e^{net_y^{(t)}}}{\sum e^{net_y^{(t)}}}$$

$$\frac{\partial E}{\partial W_{ya}} = \frac{\partial E}{\partial y} \frac{\partial y}{\partial net_y} \frac{\partial net_y}{\partial W_{ya}}$$

$$\frac{\partial E}{\partial W_{ya}} = \left[\frac{\partial E_1}{\partial y(1)} \frac{\partial y(1)}{\partial net_y^{(1)}} \frac{\partial net_y(1)}{\partial W_{ya}}\right] + \left[\frac{\partial E_2}{\partial y(2)} \cdot \frac{\partial y(2)}{\partial net_y^{(2)}} \cdot \frac{\partial net_y(2)}{\partial W_{ya}}\right] +$$

$$\left[\frac{\partial E_3}{\partial y(3)} \frac{\partial y(3)}{\partial net_y^{(3)}} \frac{\partial net_y(3)}{\partial W_{ya}}\right] + \left[\frac{\partial E_4}{\partial y(4)} \frac{\partial y(4)}{\partial net_y^{(4)}} \frac{\partial net_y(4)}{\partial W_{ya}}\right]$$

$$\boxed{\frac{\partial E_N}{\partial W_{ya}} = \sum_{t=1}^{N} \frac{\partial E_t}{\partial y(t)} \frac{\partial y(t)}{\partial net_y^{(2)}} \frac{\partial net_y(t)}{\partial W_{ya}}}$$

$$\boxed{\frac{\partial E_N}{\partial W_{ya}} = \sum_{t=1}^{N} \frac{-1}{y(t)} * y(t)(1-y(t)) * a(t)}$$

$$\boxed{\Delta W_{ya} = \eta \sum_{t=1}^{N} a(t)(1-y(t))}$$

$$\boxed{E = E_1 + E_2 + E_3 + E_4}$$

$$E = -\log\left(\frac{e^{y_1(1)}}{\sum_{i=1}^{4} e_i^{y(t)}}\right) - \log\left(\frac{e^{y_1(2)}}{\sum_{i=1}^{4} e^{y_1(2)}}\right) - \log\left(\frac{e^{y_1(3)}}{\sum e^{y_1(3)}}\right) - \log\left(\frac{e^{y(4)}}{\sum e^{y(4)}}\right)$$

(2)

$$\Delta b_y = -\eta \frac{\partial E}{\partial b_y}$$

$$\frac{\partial E}{\partial b_y} = \frac{\partial E}{\partial y} \frac{\partial y}{\partial net_y} \frac{\partial net_y}{\partial b_y}$$

$$= \left[\frac{\partial E_1}{\partial y(1)} \frac{\partial y(1)}{\partial net_y(1)} \frac{\partial net_y(1)}{\partial b_y}\right] + \left[\frac{\partial E_2}{\partial y(2)} \frac{\partial y(2)}{\partial net_y(2)} \frac{\partial net_y(2)}{\partial y}\right] +$$

$$\left[\frac{\partial E_3}{\partial y(3)} \frac{\partial y(3)}{\partial net_y(3)} \frac{\partial net_y(3)}{\partial b_y}\right] + \left[\frac{\partial E_4}{\partial y(4)} \frac{\partial y(4)}{\partial net_y(4)} \frac{\partial net_y(4)}{\partial b_y}\right]$$

$$\therefore \boxed{\frac{\partial E_N}{\partial b_y} = \sum_{t=1}^{N} \frac{\partial E_t}{\partial y(t)} \frac{\partial y(t)}{\partial net_y(t)} \frac{\partial net_y(t)}{\partial b_y}}$$

$$\boxed{\frac{\partial E_N}{\partial b_y} = \sum_{t=1}^{N} \frac{-1}{y(t)} y(t)(1-y(t))}$$

$$\frac{\partial E_N}{\partial b_y} = -\sum_{t=1}^{N} (1-y(t))$$

$$\boxed{\Delta b_y = -\eta \frac{\partial E_N}{\partial b_y} = \eta \sum_{t=1}^{N} (1-y(t))}$$

(3) $\boxed{\Delta b_a = -\eta \frac{\partial E}{\partial b_a}}$

$$\frac{\partial E}{\partial b_a} = \frac{\partial E}{\partial y} \frac{\partial y}{\partial net_y} \frac{\partial net_y}{\partial a} \frac{\partial a}{\partial b_a} \Rightarrow \boxed{\frac{\partial E_N}{\partial b_a} = \sum_{t=1}^{N} \frac{\partial E(t)}{\partial y(t)} \frac{\partial y(t)}{\partial net_y(t)} \frac{\partial net_y(t)}{\partial a(t)} \frac{\partial a(t)}{\partial b_a}}$$

$$\Rightarrow \frac{\partial E_N}{\partial b_a} = \sum_{t=1}^{N} \frac{-1}{y(t)} y(t)(1-y(t)) W_{ya} [1-(a(t))^2]$$

$$\therefore \boxed{\Delta b_a = \sum_{t=1}^{N} W_{ya} (1-y(t))[1-(a(t))^2]}$$

$$a = \tanh[W_{aa}a(t-1) + W_{ax}x(t) + b_a] \qquad \boxed{\frac{\partial}{\partial x}\tanh(x) = 1-[\tanh(x)]^2}$$

(4) $\Delta W_{aa} = -\eta \frac{\partial E_N}{\partial W_{aa}}$

$$\frac{\partial E}{\partial W_{aa}} = \frac{\partial E}{\partial y} \frac{\partial y}{\partial net_y} \frac{\partial net_y}{\partial a} \frac{\partial a}{\partial W_{aa}}$$

$$\frac{\partial E_N}{\partial W_{aa}} = \sum_{t=1}^{N} \frac{\partial E(t)}{\partial y(t)} \frac{\partial y(t)}{\partial net_y(t)} \frac{\partial net_y(t)}{\partial a(t)} \cdot \frac{\partial a(t)}{\partial W_{aa}}$$

$$= \sum_{t=1}^{N} \frac{-1}{y(t)} y(t)(1-y(t))\, W_{ya} [1-(a(t))^2]\, W_{aa}$$

$$\therefore \boxed{\Delta W_{aa} = -\eta \frac{\partial E_N}{\partial W_{aa}} = \eta \sum_{t=1}^{N} W_{ya}\, a(t) (1-y(t))[1-(a(t))^2]}$$

(5) $\Delta W_{ax} = -\eta \frac{\partial E_N}{\partial W_{ax}}$

$$\frac{\partial E_N}{\partial W_{ax}} = \sum_{t=1}^{N} \frac{\partial E(t)}{\partial y(t)} \frac{\partial y(t)}{\partial net_y(t)} \frac{\partial net_y(t)}{\partial a(t)} \frac{\partial a(t)}{\partial W_{ax}}$$

$$\frac{\partial E_N}{\partial W_{ax}} = \sum_{t=1}^{N} \frac{-1}{y(t)} y(t)(1-y(t))\, W_{ya} [1-(a(t))^2]\, a(t)$$

$$\therefore \boxed{\Delta W_{ax} = \eta \sum_{t=1}^{N} W_{ya} (1-y(t))[1-(a(t))^2]\, x(t)}$$

$$\therefore \boxed{\begin{aligned} W_{ya} &= W_{ya} + \Delta W_{ya} \\ W_{aa} &= W_{aa} + \Delta W_{aa} \\ W_{ax} &= W_{ax} + \Delta W_{ax} \\ b_y &= b_y + \Delta b_y \\ b_a &= b_a + \Delta b_a \end{aligned}}$$

→ weight updation in RNN

RNN