Backpropagation in RNN

01 December 2022 16:43

RNN intro → RNN → practical → types of RNN

→ RNN → Backprop → $\boxed{\text{BPTT}}$

Many to One RNN

Sentiment Analysis

text → 1/0

$\hat{y} \rightarrow L$

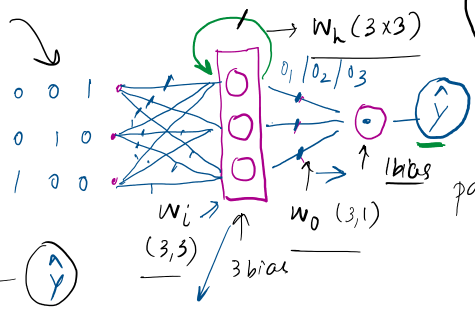
text			
cat	mat	rat	1
rat	rat	mat	1
mat	mat	cat	0

	o_1	o_2	o_3	$\hat{y}$	y
X_1	[1 0 0]	[0 1 0]	[0 0 1]		1
X_2	[0 0 1]	[0 0 1]	[0 1 0]		1
X_3	[0 1 0]	[0 1 0]	[1 0 0]		0

$\boxed{W_i \; W_h \; W_o}$

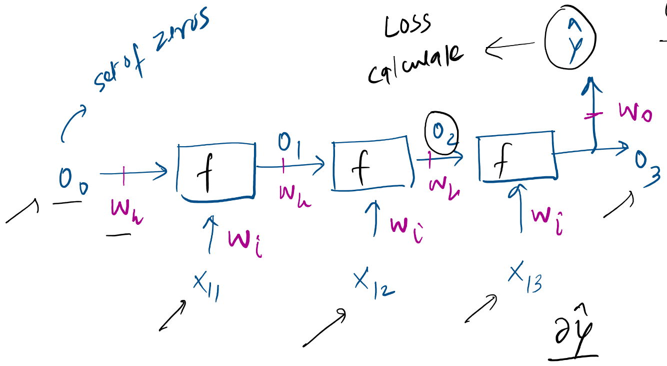
Vocab → Cat Mat Rat

[1 0 0] [0 1 0] [0 0 1]



parameters

forward prop



$$O_1 = f(X_{11}W_i + O_0W_h)$$

$$O_2 = f(X_{12}W_i + O_1W_h)$$

$$O_3 = f(X_{13}W_i + O_2W_h)$$

$$\frac{\partial \hat{y}}{\partial W_o}$$

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

y

loss calculate → minimize

$$L = -y_i \log \hat{y}_i - (1-y_i)\log(1-\hat{y}_i)$$

L min gradient descent

$$\frac{\partial L}{\partial \hat{y}}$$

W_i W_h W_o

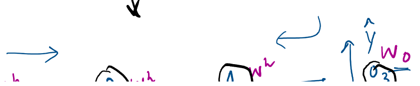
$$W_i = W_i - \eta \boxed{\frac{\partial L}{\partial W_i}}$$

$$W_h = W_h - \eta \boxed{\frac{\partial L}{\partial W_h}}$$

$$W_o = W_o - \eta \boxed{\frac{\partial L}{\partial W_o}}$$

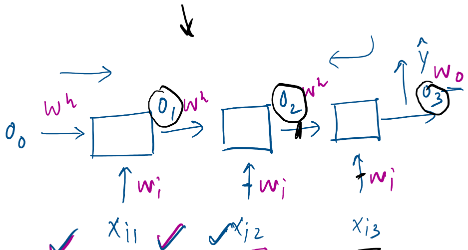
Backprop

unfold → 3 times

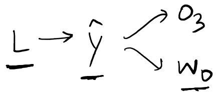


$$\frac{\partial L}{\partial W_o}$$

$L \rightarrow \hat{y} \rightarrow O_3, \; W_o$



$\frac{\partial L}{\partial W_o}$



$$\boxed{\frac{\partial L}{\partial W_o} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial W_o}}$$

$$\boxed{\frac{\partial L}{\partial W_o}}, \boxed{\frac{\partial L}{\partial W_i}}, \boxed{\frac{\partial L}{\partial W_h}}$$

$$o_1 = f(x_{i1} w_i + o_0 w_h)$$

$$o_2 = f(x_{i2} w_i + o_1 w_h)$$

$$o_3 = f(x_{i3} w_i + o_2 w_h)$$

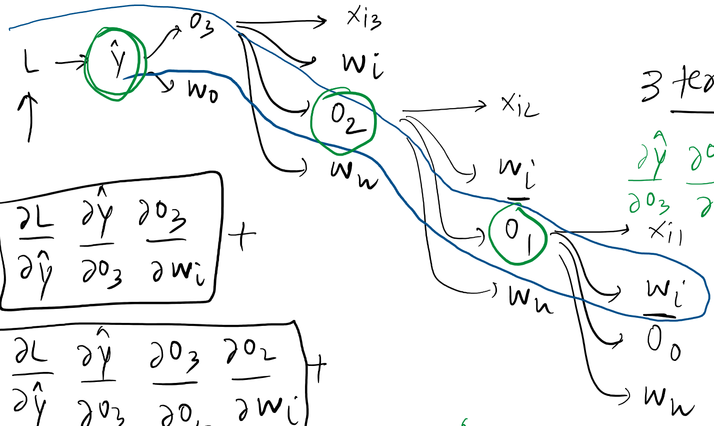
$$\hat{y} = \sigma(o_3 w_o)$$

10 words

10 terms

t=1

$\frac{\partial L}{\partial W_i}$



3 terms

$$\frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial o_2}$$

$$\boxed{\frac{\partial L}{\partial W_i}} = \boxed{\frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial W_i}} +$$

$$\boxed{\frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial o_2} \frac{\partial o_2}{\partial W_i}} +$$

$$\boxed{\frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial o_2} \frac{\partial o_2}{\partial o_1} \frac{\partial o_1}{\partial W_i}}$$

$$\frac{\partial \hat{y}}{\partial o_1} = \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial o_2} \frac{\partial o_2}{\partial o_1}$$

$$\boxed{\frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial W_i}}$$

$$\frac{\partial L}{\partial W_i} = \sum_{j=1}^{3} \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_j} \frac{\partial o_j}{\partial W_i}$$

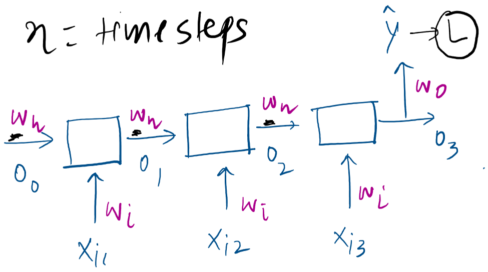
$$= \frac{\partial L}{\partial \hat{y}} \boxed{\frac{\partial \hat{y}}{\partial o_1}} \frac{\partial o_1}{\partial W_i} = \boxed{\frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial o_2} \frac{\partial o_2}{\partial o_1} \frac{\partial o_1}{\partial W_i}}$$

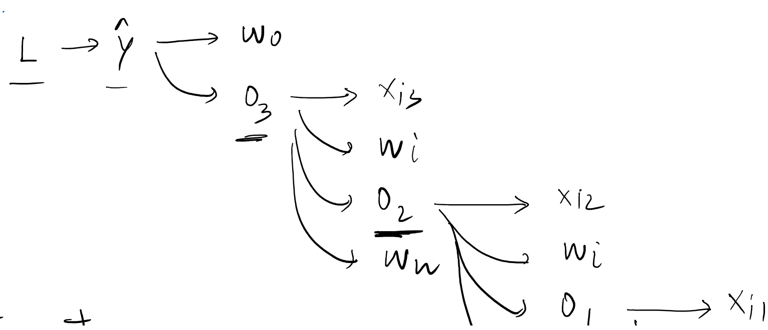
$$\frac{\partial L}{\partial \hat{y}} \boxed{\frac{\partial \hat{y}}{\partial o_2}} \frac{\partial o_2}{\partial W_i} = \boxed{\frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial o_2} \frac{\partial o_2}{\partial W_i}}$$

$$\boxed{\frac{\partial L}{\partial W_i} = \sum_{j=1}^{n} \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_j} \frac{\partial o_j}{\partial W_i}}$$

n = timesteps

$$\boxed{\frac{\partial L}{\partial W_h}}$$

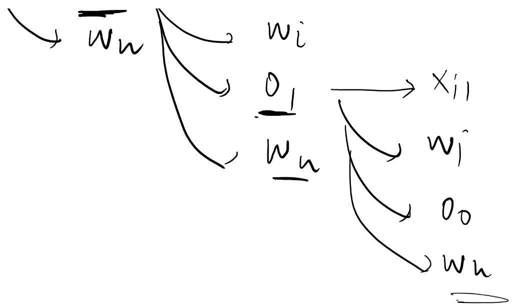




$$\frac{\partial L}{\partial W_h} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial W_h} +$$

$$\frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial o_2} \frac{\partial o_2}{\partial W_h} +$$

$$\partial W_h \quad \partial y \quad \partial o_3 \quad \partial W_h$$



$$\frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial o_2} \frac{\partial o_2}{\partial W_h} +$$

$$\frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial o_2} \frac{\partial o_2}{\partial o_1} \frac{\partial o_1}{\partial W_h}$$

$$\boxed{\frac{\partial L}{\partial W_h} = \sum_{j=1}^{n} \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_j} \frac{\partial o_j}{\partial W_h}}$$

n = timesteps

for $j=3$

$$\frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial W_h} \rightarrow \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_3} \frac{\partial o_3}{\partial o_1} \frac{\partial o_1}{\partial W_h}$$

for $j=10$

$$\frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial o_{10}} \frac{\partial o_{10}}{\partial o_1} \frac{\partial o_1}{\partial W_h} \quad \boxed{\prod_{t=2}^{j} \frac{\partial o_t}{\partial o_{t-1}}}$$

$$\prod_{t=2}^{t=3} \frac{\partial o_t}{\partial o_{t-1}} = \frac{\partial o_2}{\partial o_1} \frac{\partial o_3}{\partial o_2}$$

$$o_t = f(x_{it} W_{inp} + o_{t-1} W_h)$$

$$\frac{\partial o_t}{\partial o_{t-1}} = \prod_{t=2}^{j} f'(x_{it} W_{inp} + o_{t-1} W_h)\, W_h$$

$$\downarrow$$

$$[0-1]$$