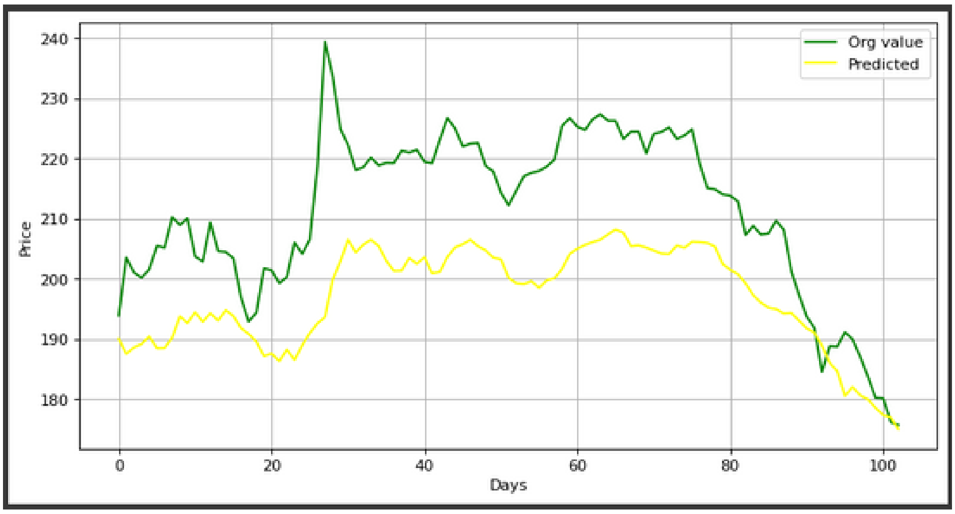
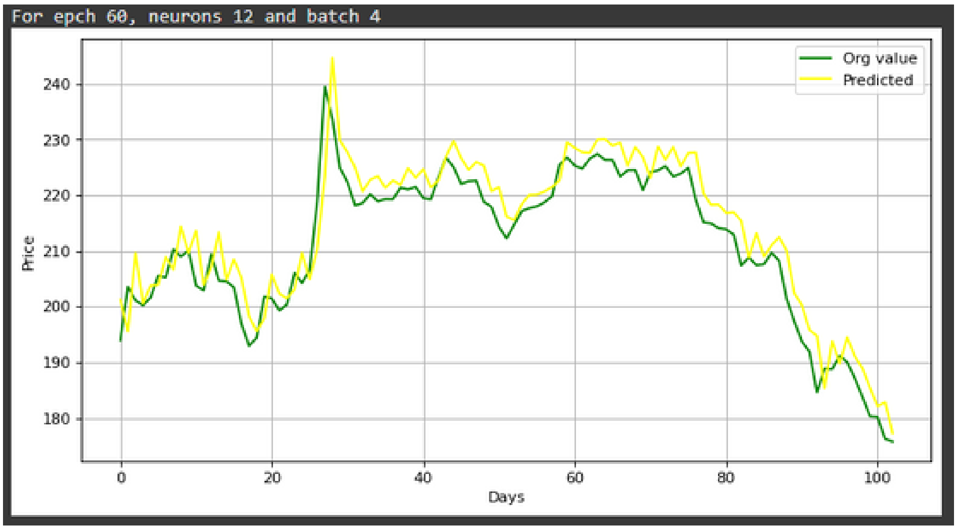
UMAIR BIN MANSOOR
Roll No : 8471
ASG - 3 (NLP)

Q 1: Comparison of RNN and LSTM



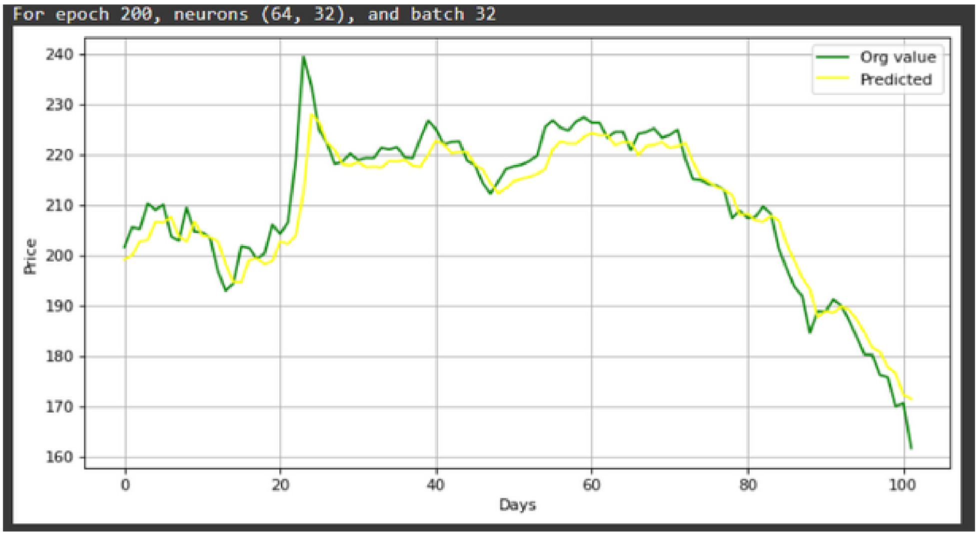
RNN - Ground truth and predicted values graph



LSTM - Ground truth and predicted values graph

The architecture of LSTM has considerably better performance than RNN. The best configuration of LSTM is 12 neurons in the hidden and dense layer, a batch size of 4 and 60 epochs.

Q 2: Comparison with Research Paper



In the research paper, Amazon stock data was taken yahoo finance API. Whereas, we have taken "EFERT" stock from PSX API.

The best performance of LSTM was at 64 and 32 neurons in the first and second layer of LSTM, with a batch size of 32 and 200 epochs.

The performance of the LSTM architecture in the research paper and in the blog is relatively the same w.r.t "EFERT" data.

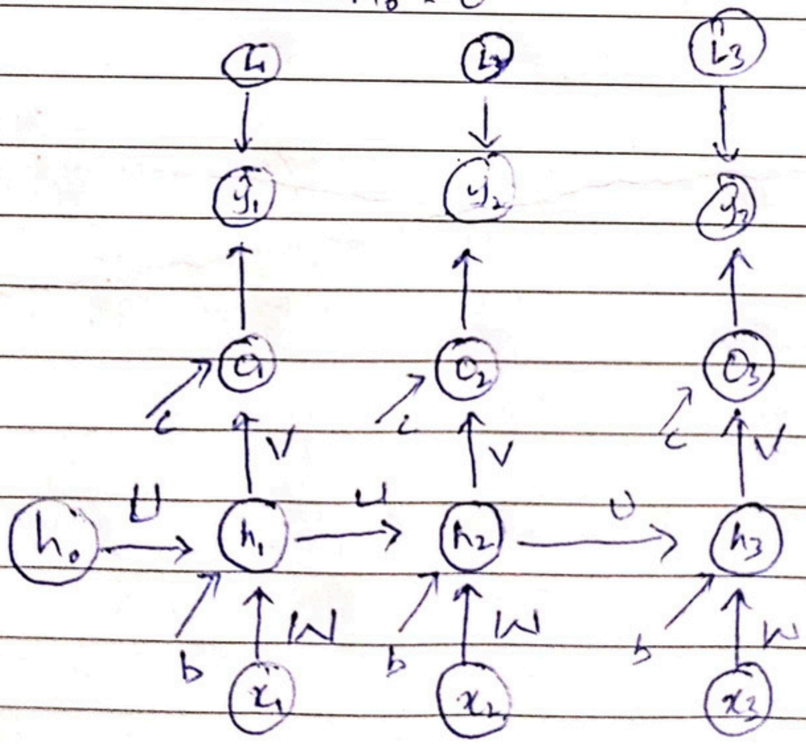
Inputs : $x_1 = 1.0$, $x_2 = 0.8$, $x_3 = 0.6$
Outputs : $y_1 = 0.5$, $y_2 = 0.6$, $y_3 = 0.7$

Parameters :

- $W = 0.5$ $(i \to h)$
- $U = 0.5$ $(h \to h)$
- $V = 0.5$ $(h \to o)$
- $b = 0.2$ (h)
- $c = 0.1$ (o)

Activation: tanh for hidden , linear for output

$h_0 = 0$



Date______20___

FORWARD PASS

Equations:

$$y_1 = Vh_1 + C$$
$$h_1 = \tanh(Wx_1 + Uh_0 + b)$$

For $t = 1$

$$y_1 = (0.5)(0.604) + 0.1 = \boxed{0.402}$$

$$h_1 = \tanh(0.5 \times 1 + 0 + 0.2) = \tanh(0.7) = 0.604$$

For $t = 2$

$$y_2 = (0.5)(0.716) + 0.1 = \boxed{0.458}$$

$$h_2 = \tanh(0.5 \times 0.8 + 0.5 \times 0.604 + 0.2) = \tanh(0.4 + 0.3 + 0.2) = 0.716$$

For $t = 3$

$$y_3 = (0.5)(0.858) + 0.1 = \boxed{0.529}$$

$$h_3 = \tanh(0.5 \times 0.6 + 0.5 \times 0.716 + 0.2) = \tanh(0.3 + 0.358 + 0.2) = 0.858$$

Loss (MSE):

$$L_1 = \frac{1}{2}(y_1 - \hat{y}_1)^2 = \frac{1}{2}(0.5 - 0.402)^2 = 4.8 \times 10^{-3}$$

$$L_2 = \frac{1}{2}(0.6 - 0.458)^2 = 0.01$$

$$L_3 = \frac{1}{2}(0.7 - 0.529)^2 = 0.01$$

$$L = L_1 + L_2 + L_3 = \boxed{0.024}$$

BACK PROPAGATION

PARAMS : V, U, W, b, c

At time t = 3

$$\hat{y}_3 = Vh_3 + c$$

$$h_3 = \tanh(Wx_3 + Uh_2 + b)$$

First cal for V: Because V has no dependencies on previous hidden states

$$\frac{\partial L_3}{\partial V} = \frac{\partial L_3}{\partial y_3} \frac{\partial y_3}{\partial V} \cancel{\frac{\partial y_3}{\partial x}}$$

$$\frac{\partial}{\partial V}(Vh_3 + c)$$

$$\frac{\partial y_3}{\partial V} = h_3 = \boxed{0.858}$$

$$\frac{\partial L_3}{\partial y_3} = \frac{\partial}{\partial \hat{y}_3} \frac{1}{2}(y - \hat{y})^2$$

$$= (y - \hat{y})(-1)$$

$$= \hat{y} - y$$

$$= 0.529 - 0.7$$

$$\frac{\partial L_3}{\partial y_3} = \boxed{-0.171}$$

$$\frac{\partial L_3}{\partial V} = 0.858 \times -0.171$$

$$= -0.146$$

$$\boxed{\frac{\partial L_3}{\partial V} = -0.146}$$

Date______20___

Cal For c

$$\frac{\partial L_3}{\partial c} = \frac{\partial L_3}{\partial y_3}\frac{\partial y_3}{\partial c} = (-0.171)\left(\frac{\partial y_3}{\partial c}\right)$$

Now,

$$\frac{\partial y_3}{\partial c} = \frac{\partial}{\partial c}(Vh_3 + c) = 1$$

$$\boxed{\frac{\partial L_3}{\partial c} = -0.171}$$

Cal For U

① — $\frac{\partial L_3}{\partial U} = \frac{\partial L_3}{\partial y_3}\frac{\partial y_3}{\partial h_3}\frac{\partial h_3}{\partial U}$

$\frac{\partial L_3}{\partial y_3} = -0.171$

$\frac{\partial y_3}{\partial h_3} = V = 0.5$

② - $\frac{\partial h_3}{\partial U} = \frac{\partial h_3}{\partial u} + \frac{\partial h_3}{\partial h_2}\frac{\partial h_2}{\partial u}$

$= \cancel{\frac{\partial h_3}{\partial u} + \frac{\partial h_3}{\partial h_2}} [$

$\frac{\partial h_2}{\partial u} = \frac{\partial h_2}{\partial u} + \frac{\partial h_2}{\partial h_1}\frac{\partial h_1}{\partial u} \longrightarrow \frac{\partial}{\partial u}(\tanh(wx_1 + uh_0 + b))$

$= \operatorname{sech}^2(wx_1 + uh_0 + b)\, h_0$

$\frac{\partial h_1}{\partial u} = 0$

$\frac{\partial h_2}{\partial U} = \operatorname{sech}^2(0.4 + 0.3 + 0.1)(0.604)$

$= (0.55)(0.604)$

$\frac{\partial h_2}{\partial u} = \boxed{0.337}$

ⓐ

$\frac{\partial h_2}{\partial u} = \operatorname{sech}^2(wx_2 + uh_1 + c)\, h_1$

$= \operatorname{sech}^2(0.5 \times 0.8 + 0.5 \times 0.6 + 0.1)(0.604)$

QUALITY PAPER PRODUCTS

CamScanner

Date________ 20___

$$\frac{\partial h_3}{\partial h_2} = \frac{\partial}{\partial h_2}\left(\tanh(\cancel{W}x_3 + Uh_2 + b)\right)$$

$$= \operatorname{sech}^2(wx_3 + Uh_2 + b) \times U$$

$$= \operatorname{sech}^2(0.5 \times 0.6 + 0.5 \times 0.716 + 0.2) \times 0.5$$

$$= \operatorname{sech}^2(0.3 + 0.358 + 0.2) \times 0.5$$

$$= \operatorname{sech}^2(0.758) \times 0.5$$

$$= 0.59 \times 0.5$$

$$\boxed{\frac{\partial h_3}{\partial h_2} = 0.29} \quad \text{— b}$$

$$\frac{\partial h_3}{\partial u} = \frac{\partial}{\partial u}\left(\tanh(wx_3 + Uh_2 + b)\right)$$

$$= \operatorname{sech}^2(wx_3 + uh_2 + c) \times h_2$$

$$= 0.59 \times 0.716$$

$$\boxed{\frac{\partial h_3}{\partial u} = 0.422} \quad \text{— c}$$

Now, $\frac{\partial h_3}{\partial U} = \frac{\partial h_3}{\partial u} + \frac{\partial h_3}{\partial h_2}\frac{\partial h_2}{\partial u}$

Putting values a, b, c

$$= 0.422 + (0.29)(0.337)$$

$$= 0.422 + 0.09$$

$$\boxed{\frac{\partial h_3}{\partial U} = 0.51}$$

Now,

$$\frac{\partial L_3}{\partial U} = \frac{\partial L_3}{\partial y_3}\frac{\partial y_3}{\partial h_3}\frac{\partial h_3}{\partial U}$$

$$= (-0.171)(0.5)(0.51) = -0.04$$

$$\boxed{\frac{\partial L_3}{\partial U} = -0.04}$$

$\boxed{\text{Cal For } W}$

$$\frac{\partial L_3}{\partial W} = \frac{\partial L_3}{\partial y_3}\frac{\partial y_3}{\partial h_3}\frac{\partial h_3}{\partial W}$$

$$\frac{\partial h_3}{\partial W} = \frac{\partial h_3}{\partial W} + \frac{\partial h_3}{\partial h_2}\frac{\partial h_2}{\partial W}$$

$$\frac{\partial h_2}{\partial W} = \frac{\partial h_2}{\partial W} + \frac{\partial h_2}{\partial h_1}\frac{\partial h_1}{\partial W}$$

$$\frac{\partial h_1}{\partial W} = \frac{\partial}{\partial W}(\tanh(wx_1 + uh_0 + b))$$

$$= \operatorname{sech}^2(wx_1 + uh_0 + b)x_1$$

$$= \operatorname{sech}^2(0.5 + 0 + 0.1)(1)$$

$$\boxed{\frac{\partial h_1}{\partial W} = 0.71}$$

$$\frac{\partial h_2}{\partial h_1} = \frac{\partial}{\partial h_1}(\tanh(wx_2 + uh_1 + b))$$

$$= \operatorname{sech}^2(wx_2 + uh_1 + c)\cdot u$$

$$\frac{\partial h_2}{\partial h_1} = \operatorname{sech}^2(0.4 + 0.3 + 0.1)\cdot 0.5 \boxed{= 0.27}$$

QUALITY PAPER PRODUCTS

CamScanner

$$\frac{\partial h^2}{\partial W} = \frac{\partial}{\partial W}\left(\tanh\left(Wx_2 + uh_1 + b\right)\right)$$

$$= \operatorname{sech}^2\left(Wx_2 + uh_1 + b\right) x_2$$

$$= \operatorname{sech}^2\left(0.4 + 0.3 + 0.2\right) \times 0.8$$

$$= 0.55 \times 0.8 = \boxed{0.44}$$

$$\frac{\partial h^2}{\partial W} = 0.44 + 0.27 \times 0.71 = \boxed{0.63}$$

$$\frac{\partial h_3}{\partial h_2} = 0.29$$

$$\frac{\partial h_3}{\partial W} = \frac{\partial}{\partial W}\left(\tanh\left(Wx_3 + uh_2 + b\right)\right)$$

$$= 0.59 \times 0.6 = \boxed{0.354}$$

$$\frac{\partial h_3}{\partial W} = \frac{\partial h_3}{\partial W} + \frac{\partial h_3}{\partial h_2}\frac{\partial h_2}{\partial W}$$

$$\boxed{\frac{\partial h_3}{\partial W} = 0.354 + 0.29 \times 0.63 = 0.536}$$

$$\frac{\partial L_3}{\partial W} = \frac{\partial L_3}{\partial y_3}\frac{\partial y_3}{\partial h_3}\frac{\partial h_3}{\partial W}$$

$$= (-0.171)(0.5)(0.536)$$

$$\boxed{\frac{\partial L_3}{\partial W} = -0.04}$$

QUALITY PAPER PRODUCTS

CamScanner

$\boxed{\text{Cal for } b}$

$$\frac{\partial L_3}{\partial b} = \frac{\partial L_3}{\partial y_3}\frac{\partial y_3}{\partial h_3}\frac{\partial h_3}{\partial b}$$

$$\frac{\partial h_3}{\partial b} = \frac{\partial h_3}{\partial b} + \frac{\partial h_3}{\partial h_2}\frac{\partial h_2}{\partial b}$$

$$\frac{\partial h_2}{\partial b} = \frac{\partial h_2}{\partial b} + \frac{\partial h_2}{\partial h_1}\frac{\partial h_1}{\partial b}$$

$$\frac{\partial h_1}{\partial b} = \frac{\partial}{\partial b}(\tanh(wx_1 + uh_0 + b))$$

$$= \operatorname{sech}^2(wx_1 + uh_0 + b)\cdot 1 = \boxed{0.71}$$

$$\frac{\partial h_2}{\partial h_1} = 0.27$$

$$\frac{\partial h_2}{\partial b} = \operatorname{sech}^2(wx_2 + uh_1 + b)1 = \boxed{0.55}$$

$$\frac{\partial h_2}{\partial b} = 0.55 + 0.27 \times 0.71 = 0.55 + 0.19$$

$$\boxed{\frac{\partial h_2}{\partial b} = 0.74}$$

$$\boxed{\frac{\partial h_3}{\partial h_2} = 0.29}$$

$$\frac{\partial h_3}{\partial b} = \operatorname{sech}^2(wx_3 + uh_2 + b)\ 1 = 0.59$$

$$\boxed{\frac{\partial h_3}{\partial b} = 0.59}$$

QUALITY PAPER PRODUCTS

CamScanner

$$\frac{\partial h_3}{\partial b} = 0.59 + (0.29)(0.74)$$

$$= 0.59 + 0.21 = \boxed{0.8 = \frac{\partial h_3}{\partial b}}$$

$$\frac{\partial L_3}{\partial b} = \frac{\partial L_3}{\partial y_3}\frac{\partial y_3}{\partial h_3}\frac{\partial h_3}{\partial b}$$

$$= (-0.171)(0.5)(0.8)$$

$$\boxed{\frac{\partial L_3}{\partial b} = -0.06}$$

Gradiant of

So, / Losses at $t = 3$:

$$\frac{\partial L_3}{\partial V} = -0.146$$

$$\frac{\partial L_3}{\partial c} = -0.171$$

$$\frac{\partial L_3}{\partial U} = -0.04$$

$$\frac{\partial L_3}{\partial W} = -0.04$$

$$\frac{\partial L_3}{\partial b} = -0.06$$

CamScanner

Date______20___

FOR TIME STEP : $T = 2$

$$y_2 = Vh_2 + c$$
$$h_2 = \tanh(Wx_2 + Uh_1 + b)$$

First, cal for V & C

For V

$$\frac{\partial L_2}{\partial V} = \frac{\partial L_2}{\partial y_2}\frac{\partial y_2}{\partial V}$$

$$\frac{\partial y_2}{\partial V} = \frac{\partial}{\partial V}(Vh_2 + c) = h_2 = 0.51$$

$$\frac{\partial L_2}{\partial \hat{y}_2} = \frac{\partial}{\partial \hat{y}_2}\left[\frac{1}{2}(y - \hat{y})^2\right] = \hat{y}_2 - y_2 = 0.458 - 0.6 = -0.14$$

$$\boxed{\frac{\partial L_2}{\partial V} = (0.51)(-0.14) = -0.07}$$

For C

$$\frac{\partial L_2}{\partial c} = \frac{\partial L_2}{\partial y_2}\frac{\partial y_2}{\partial c}$$

$$\frac{\partial y_2}{\partial c} = \frac{\partial}{\partial c}(Vh_2 + c) = 1$$

$$\boxed{\frac{\partial L_2}{\partial c} = (-0.14)(1) = -0.14}$$

CamScanner

$\boxed{\text{For } U}$

$$\frac{\partial L_2}{\partial u} = \frac{\partial L_2}{\partial y_2}\frac{\partial y_2}{\partial h_2}\frac{\partial h_2}{\partial u}$$

$$\frac{\partial h_2}{\partial u} = \frac{\partial h_2}{\partial u} + \frac{\partial h_2}{\partial h_1}\frac{\partial h_1}{\partial u} = 0.337 \quad \text{(From prev calculations)}$$

$$\frac{\partial}{\partial h_2}(Vh_2 + c) = V = \boxed{0.5 = \frac{\partial y_2}{\partial h_2}}$$

$$\frac{\partial L_2}{\partial y_2} = -0.14 \quad \text{(From previous calculations)}$$

So,

$$\frac{\partial L_2}{\partial u} = (-0.14)(0.5)(0.33) = \boxed{-0.23}$$

$\boxed{\text{For } W}$

$$\frac{\partial L_2}{\partial W} = \frac{\partial L_2}{\partial y_2}\frac{\partial y_2}{\partial h_2}\frac{\partial h_2}{\partial W}$$

$$\frac{\partial h_2}{\partial W} = \frac{\partial h_2}{\partial W} + \frac{\partial h_2}{\partial h_1}\frac{\partial h_1}{\partial W} = 0.63 \quad \text{(From prev cal)}$$

$$\frac{\partial L_2}{\partial W} = (-0.14)(0.5)(0.63)$$

$$\boxed{\frac{\partial L_2}{\partial W} = -0.04}$$

QUALITY PAPER PRODUCTS

CamScanner

$\boxed{\text{For } b}$

$$\frac{\partial L_2}{\partial b} = \frac{\partial L_2}{\partial y_2} \frac{\partial y_2}{\partial h_2} \frac{\partial h_2}{\partial b}$$

$$\frac{\partial h_2}{\partial b} = \frac{\partial h_2}{\partial b} + \frac{\partial h_2}{\partial h_1} \frac{\partial h_1}{\partial b} = 0.74 \quad (\text{Prev cal})$$

$$\frac{\partial L_2}{\partial b} = (-0.14)(0.5)(0.74) = \boxed{-0.05}$$

All gradiants of Losses :

$$\frac{\partial L_2}{\partial V} = -0.07$$

$$\frac{\partial L_2}{\partial c} = -0.14$$

$$\frac{\partial L_2}{\partial U} = -0.23$$

$$\frac{\partial L_2}{\partial W} = -0.04$$

$$\frac{\partial L_2}{\partial b} = -0.05$$

CS CamScanner

Date______ 20___

$$\boxed{\text{For TIME STEP : } T = 1}$$

$$y_1 = Vh_1 + c$$
$$h_1 = \tanh(Wx_1 + Uh_0 + b)$$

$\boxed{\text{For } c}$

$$\frac{\partial L_1}{\partial c} = \frac{\partial L_1}{\partial y_1} \frac{\partial y_1}{\partial c}$$

$$\frac{\partial L_1}{\partial y_1} = \frac{\partial}{\partial y_1}\left[\frac{1}{2}(y - \hat{y})^2\right] = (\hat{y}_1 - y_1) = 0.402 - 0.5$$
$$= \boxed{-0.102}$$

$$\frac{\partial y_1}{\partial c} = \frac{\partial}{\partial c}(Vh_1 + c) = 1$$

$$\boxed{\frac{\partial L_1}{\partial c} = -0.102}$$

$\boxed{\text{For } V}$

$$\frac{\partial L_1}{\partial V} = \frac{\partial L_1}{\partial y_1} \frac{\partial y_1}{\partial V}$$

$$\frac{\partial y_1}{\partial V} = \frac{\partial}{\partial V}(Vh_1 + c) = h_1 = 0.604$$

$$\boxed{\frac{\partial L_1}{\partial V} = (-0.102)(0.604) = -0.06}$$

$\boxed{\text{for } u}$

$$\frac{\partial L_1}{\partial u} = \frac{\partial L_1}{\partial y_1}\frac{\partial y_1}{\partial h_1}\frac{\partial h_1}{\partial u}$$

$$\frac{\partial h_1}{\partial u} = \frac{\partial}{\partial u}\left(\tanh(Wx_1 + uh_0 + b)\right)$$

$$= \text{sech}^2(Wx_1 + uh_0 + b)\, h_0$$

$$= 0 \rightarrow \text{since } h_0 = 0$$

$$\frac{\partial L_1}{\partial u} = 0$$

$\boxed{\text{for } W}$

$$\frac{\partial L_1}{\partial W} = \frac{\partial L_1}{\partial y_1}\frac{\partial y_1}{\partial h_1}\frac{\partial h_1}{\partial W}$$

$$\frac{\partial}{\partial W}\left(\tanh(wx_1 + uh_0 + b)\right)$$

$$= \text{sech}^2(Wx_1 + uh_0 + b)\, x_1$$

$$\boxed{= 0.71}$$

$$\frac{\partial}{\partial h_1}(Vh_1 + c) = V = \boxed{0.5}$$

$$\frac{\partial L_1}{\partial W} = (-0.102)(0.5)(0.71)\ \boxed{= -0.036}$$

QUALITY PAPER PRODUCTS

CS CamScanner

$\boxed{\text{For } b}$

$$\frac{\partial L_1}{\partial b} = \frac{\partial L_1}{\partial y_1}\frac{\partial y_1}{\partial h_1}\frac{\partial h_1}{\partial b} = -0.102 \times 0.5 \times 0.71$$

$$\frac{\partial L_1}{\partial b} = -0.036$$

So now, taking Sum of all Gradients.

$$\frac{\partial L}{\partial c} = -0.102 - 0.14 - 0.171 = -0.413$$

$$\frac{\partial L}{\partial u} = -0.04 - 0.23 + 0 = -0.27$$

$$\frac{\partial L}{\partial w} = -0.036 - 0.04 - 0.04 = -0.116$$

$$\frac{\partial L}{\partial v} = -0.146 - 0.07 - 0.06 = -0.276$$

$$\frac{\partial L}{\partial b} = -0.036 - 0.05 - 0.06 = -0.146$$

CS CamScanner

Date ______ 20___

Updating Weights of Parameters.

Considering $\eta = 0.1$

$$C_2 = C_1 - \eta \frac{\partial L}{\partial c} = 0.1 - (0.1)(-0.413)$$

$$= 0.1 + 0.041 = \boxed{0.141}$$

$$b_2 = b_1 - \eta \frac{\partial L}{\partial b} = 0.2 - (0.1)(-0.146)$$

$$= 0.2 + 0.014 = \boxed{0.214}$$

$$V_2 = V_1 - \eta \frac{\partial L}{\partial V} = 0.5 - (0.1)(-0.276)$$

$$= 0.5 + 0.027 = \boxed{0.527}$$

$$W_2 = W_1 - \eta \frac{\partial L}{\partial W} = 0.5 - (0.1)(-0.116)$$

$$= 0.5 + 0.011 = \boxed{0.511}$$

$$U_2 = U_1 - \eta \frac{\partial L}{\partial U} = 0.5 - (0.1)(-0.27)$$

$$= 0.5 + 0.027 = \boxed{0.527}$$

CamScanner