CS 271 Spring 2020
Assignment 6

Student ID: 014498887
Name: Yash Sahasrabuddhe

Q.1.

Q.1.

$V_0 = x$

$V_1 = y$

$V_2 = -(V_0 + V_1)$

$V_3 = 1 + e^{V_2}$

$V_4 = V_0 / V_3$

$Z = V_4$

From above equations,

$\Rightarrow \frac{dV_2}{dV_0} = -1 \quad \& \quad \frac{dV_2}{dV_1} = -1$

$\frac{dV_3}{dV_2} = e^{V_2}$

$\frac{dV_4}{dV_0} = \frac{1}{V_3} \qquad \frac{dV_4}{dV_3} = \frac{-V_0}{V_3^2}$

$\frac{dz}{dV_3} = \frac{dz}{dV_4} \cdot \frac{dV_4}{dV_3} = \frac{dz}{dV_4} \cdot \frac{-V_0}{V_3^2}$

$\frac{dz}{dV_2} = \frac{dz}{dV_3} \cdot \frac{dV_3}{dV_2} = e^{V_2} \cdot \frac{dz}{dV_3}$

$\frac{dz}{dV_1} = \frac{dz}{dV_2} \cdot \frac{dV_2}{dV_1} = -1 \cdot \frac{dz}{dV_2}$

$\frac{dz}{dV_0} = \frac{dz}{dV_2} \cdot \frac{dV_2}{dV_0} + \frac{dz}{dV_4} \cdot \frac{dV_4}{dV_0} = -1\frac{dz}{dV_2} + \frac{1}{V_3}\frac{dz}{dV_4}$

Therefore, by backward pass,

$$dz = 1$$
$$dv_4 = dz$$
$$dv_3 = \left(-v_0 / v_3^2\right) dv_4$$
$$dv_2 = e^{v_2} dv_3$$
$$dv_1 = -dv_2$$
$$dv_0 = -dv_2 + 1/v_3 \cdot dv_4$$

Q.2.

a)

Q.2. (a)

$$Y = W_4 \, f(W_0 X_0 , W_1 X_1) + W_5 \, f(W_1 X_0 , W_3 X_1)$$

$$\Rightarrow Y = W_4 \cdot \max(W_0 X_0 + W_2 X_1 , 0) + W_5 \cdot \max(W_1 X_0 + W_3 X_1 , 0)$$

b)

(b) $W_0 = 1$ $\quad W_2 = -1$ $\quad W_4 = 1$

$W_1 = -1$ $\quad W_3 \;\; 1$ $\quad W_5 = 1$

$\therefore$ According to the function in (a),

If $X_0 = 0$ & $X_1 = 0$,

$$Y = 1\max(1(0) + -1(0)) + 1\max(-1(0) + (1)(0))$$

$$\boxed{Y = 0} \quad -(1)$$

If $X_0 = 0$ & $X_1 = 1$, then,

$$Y = 1.\max((1)(0) + (-1)(1), 0) + 1.\max((-1)(0) + (1)(1), 0) = 1$$

$$\Rightarrow \boxed{Y = 1} \quad -(2)$$

If $X_0 = 1$ & $X_1 = 0$

$$Y = 1\max((1)(1) + (-1)(0), 0) + 1.\max((-1)(1) + (1)(0), 0)$$

$$\Rightarrow \boxed{Y = 1} \quad -(3)$$

Finally, $X_0 = 1$ & $X_1 = 1$,

$$Y = 1.\max(((1)(1) + (-1)(1), 0) + 1.\max((-1)(1) + (1)(1), 0)$$

$$\Rightarrow \boxed{Y = 0} \quad -(4)$$

$\therefore$ Truth Table

X_0	X_1	Y	
0	0	0	= (1)
0	1	1	= (2)
1	0	1	= (3)
1	1	0	= (4)

As all eq^n are equal.

Hence Proved.

Q.4.

Q.4. Given:

$$f(x_0, x_1) = a x_0 + b x_1$$

$$g(s, t) = s + t$$

From MLP equation, we have,

$$g(s,t) = w_4 (a w_0 x_0 + b w_2 x_1) + w_5 (a w_0 x_0 + b w_2 x_1) \quad -①$$

Since, this equation is linear and represents the form of a single layer perceptron:

$$f(x,y) = w_0 x + w_1 y + b. \quad -②$$

As eq^n ① & eq^n ② are analogous, it shows that this is equivalent to single layer perceptron.

Eq^n ① represents the perceptron.

Q.6.

a) The output for this part is as below. The source code is included in the submission as Q6a.py file.

```
/Users/yash/PycharmProjects/ML/venv/bin/python /Users/yash/PycharmProjects/ML/Q6a.py
The updated weights are:
[1.316711015500443, 4.78459899823054, 0.2835419844736556, 2.32096114996884, -4.4504230640874045, 3.8446564337384994]
('Y for X0_test ', 0.55, 'and for X1_test ', 0.11, ' is: ', 0.6137817719345118)
('Y for X0_test ', 0.32, 'and for X1_test ', 0.21, ' is: ', 0.6435909474249732)
('Y for X0_test ', 0.24, 'and for X1_test ', 0.64, ' is: ', 0.8194976023374765)
('Y for X0_test ', 0.86, 'and for X1_test ', 0.68, ' is: ', 0.315723406584258)
('Y for X0_test ', 0.53, 'and for X1_test ', 0.79, ' is: ', 0.6126744690754964)
('Y for X0_test ', 0.46, 'and for X1_test ', 0.54, ' is: ', 0.6956929094559849)
('Y for X0_test ', 0.16, 'and for X1_test ', 0.51, ' is: ', 0.7490747701240221)
('Y for X0_test ', 0.52, 'and for X1_test ', 0.94, ' is: ', 0.5985885405092732)
('Y for X0_test ', 0.46, 'and for X1_test ', 0.87, ' is: ', 0.6689417453121891)
('Y for X0_test ', 0.96, 'and for X1_test ', 0.63, ' is: ', 0.23585478985702268)

Predicted Values for Test Case:
[1, 1, 1, 0, 1, 1, 1, 1, 1, 0]

Actual Values for Test Case
[1, 0, 1, 0, 0, 1, 1, 0, 1, 0]
('Accuracy is:', 70.0)

Process finished with exit code 0
```

b) The output for this part is as below. The source code is included in the submission as Q6b.py file.

```
/Users/yash/PycharmProjects/ML/venv/bin/python /Users/yash/PycharmProjects/ML/Q6b.py
The updated weights are:
[0.9682395212755759, 4.688864440802466, 0.6282056770325203, 2.622787374148792, -18.73657201358637, 14.376878853979774]
('Y for X0_test ', 0.55, 'and for X1_test ', 0.11, ' is: ', 1.499160156675794)
('Y for X0_test ', 0.32, 'and for X1_test ', 0.21, ' is: ', 1.334391384886155)
('Y for X0_test ', 0.24, 'and for X1_test ', 0.64, ' is: ', 1.311549616844328)
('Y for X0_test ', 0.86, 'and for X1_test ', 0.68, ' is: ', -0.2616793144820466)
('Y for X0_test ', 0.53, 'and for X1_test ', 0.79, ' is: ', 0.495330893520908)
('Y for X0_test ', 0.46, 'and for X1_test ', 0.54, ' is: ', 1.1184475705184163)
('Y for X0_test ', 0.16, 'and for X1_test ', 0.51, ' is: ', 1.2376483030808316)
('Y for X0_test ', 0.52, 'and for X1_test ', 0.94, ' is: ', 0.23479430665482148)
('Y for X0_test ', 0.46, 'and for X1_test ', 0.87, ' is: ', 0.5413076267187638)
('Y for X0_test ', 0.96, 'and for X1_test ', 0.63, ' is: ', -0.45660936111059414)

Predicted Values for Test Case:
[1, 1, 1, 0, 0, 1, 1, 0, 1, 0]

Actual Values for Test Case
[1, 0, 1, 0, 0, 1, 1, 0, 1, 0]
('Accuracy is:', 90.0)

Process finished with exit code 0
```

Q.8.

Q.8. For mini-batch, we will take four inputs, X_0, X_1, X_2, X_3.

Then, error function becomes,

$$E(w) = \frac{1}{2}\left(\frac{w_4}{1+e^{-(w_0 X_0 + w_2 X_1)}} + \frac{w_5}{1+e^{-(w_1 X_0 + w_3 X_1)}} - Z\right)^2$$

$$+ \frac{1}{2}\left(\frac{w_4}{1+e^{-(w_0 X_2 + w_2 X_3)}} + \frac{W_5}{1+e^{-(w_0 X_2 + w_2 X_3)}} - Z\right)^2$$

Forward Pass:

1. $V_0 = W_0$
2. $V_1 = W_1$
3. $V_2 = W_2$
4. $V_3 = W_3$
5. $V_4 = W_4$
6. $V_5 = W_5$
7. $V_6 = X_0 V_0 + X_1 V_2$
8. $V_7 = X_0 V_1 + X_1 V_3$
9. $V_8 = 1 + e^{-V_6}$
10. $V_9 = 1 + e^{-V_7}$
11. $V_{10} = V_4 / V_8$
12. $V_{11} = V_5 / V_9$
13. $V_{12} = (V_{10} + V_{11} - Z)^2 / 2$
14. $V_{13} = X_2 V_0 + X_3 V_2$
15. $V_{14} = X_2 V_1 + X_3 V_3$
16. $V_{15} = 1 + e^{-V_{13}}$
17. $V_{16} = 1 + e^{-V_{14}}$
18. $V_{17} = V_4 / V_{15}$
19. $V_{18} = V_5 / V_{16}$

20. $v_{19} = (v_{17} + v_{18} - z)^2 / 2$

21. $v_{20} = v_{12} + v_{19}$

22. $z = v_{20}$

Backward Pass:

1. $dz = 1$
2. $dv_{18} = v_{17} + v_{18} - z$
3. $dv_{17} = v_{17} + v_{18} - z$
4. $dv_{16} = -v_5 / v_{16}^2 \, dv_{18}$
5. $dv_{15} = -v_4 / v_{15}^2 \, dv_{17}$
6. $dv_{14} = -e^{-v_{14}} \, dv_{16}$
7. $dv_{13} = -e^{-v_{13}} \, dv_{15}$
8. $dv_{11} = v_{10} + v_{11} - z$
9. $dv_{10} = v_{10} + v_{11} - z$
10. $dv_9 = -v_5 / v_9^2 \, dv_{11}$
11. $dv_8 = -v_4 / v_8^2 \, dv_{10}$
12. $dv_7 = -e^{-v_7} \, dv_9$
13. $dv_6 = -e^{v_6} \, dv_8$
14. $dv_5 = dv_{11} / v_9 + dv_{18} / v_{16}$
15. $dv_4 = dv_{10} / v_8 + dv_{17} / v_{15}$
16. $dv_3 = x_1 \, dv_7 + x_3 \, dv_{14}$
17. $dv_2 = x_1 \, dv_6 + x_3 \, dv_{13}$
18. $dv_1 = x_0 \, dv_7 + x_2 \, dv_{14}$
19. $dv_0 = x_0 \, dv_7 + x_2 \, dv_{13}$

Q.10.

I have used the HMM_fast_ref code provided by the professor for this question. The files for this code can be found in the submission. The main source code is hmm.c. The final output which I got is as below:

```
T = 50000, N = 2, M = 27, iterations = 700

final pi =
 0.47879  0.52121 ,  sum = 1.000000

final A =
 0.34551  0.65449 ,  sum = 1.000000
 0.00000  1.00000 ,  sum = 1.000000

final B^T =
a  0.03471  0.06191
b  0.03470  0.01577
c  0.03473  0.02633
d  0.03471  0.03118
e  0.03471  0.09971
f  0.03473  0.01930
g  0.03472  0.01704
h  0.03471  0.03285
i  0.03472  0.05376
j  0.03471  0.01114
k  0.03470  0.01230
l  0.03470  0.03312
m  0.03471  0.02039
n  0.03470  0.05209
o  0.03470  0.05841
p  0.03470  0.01972
q  0.03471  0.01077
r  0.03472  0.04574
s  0.03471  0.05032
t  0.03470  0.07249
u  0.03470  0.02154
v  0.03469  0.01399
w  0.03472  0.01574
x  0.03521  0.01137
y  0.03767  0.01641
z  0.03482  0.01068
   0.09400  0.16592
sum[0] = 1.000000 sum[1] = 1.000000

log [P(observations | lambda)] = -144152.892981

(base) Yashs-MBP:Assignment6 yash$
```

Note – I discussed the questions with Aditi Walia. No code was shared.