Name – Ahaan Tagare
Student ID – 33865799

Coursework 1 Part A

1.a – Calculations to process the first training example
(x1 = 0, x2 = 1, target = 1)
w1: -0.1, w2: -0.3, w3: -0.2, w4: 0.2, w5: 0.1, w6: -0.1, w7: 0.2, w8: -0.3, w9: 0.3

Weights w 10 and 11 are taken as 0 to connect x1 to sigmoid 3 and and x2 to sigmoid 1 to create a fully connected neural network.

1.a.i – Computed outputs for each network node during the forward pass (4 marks)

1. Sigmoid 1 (y1):
   net1 = w1 * x1 + w11 * x2 = -0.1 * 0 + 0 * 1 = 0
   y1 = 1 / (1 + e^(-0)) = 0.5000
2. Sigmoid 2 (y2):
   net2 = w3 * x1 + w4 * x2 = -0.2 * 0 + 0.2 * 1 = 0.2000
   y2 = 1 / (1 + e^(-0.2)) ≈ 0.5498
3. Sigmoid 3 (y3):
   net3 = w6 * x2 + w10 * x1 = -0.1 * 1 + 0 * 0 = -0.1000
   y3 = 1 / (1 + e^(0.1)) ≈ 0.4750
4. Sum Node (output):
   output = w2*x1* + *w5*x2 + w9*y1* + *w8*y2 + w7*y3*
   = *-0.3*0 + 0.1*1* + *0.3*0.5000 + (-0.3)*0.5498* + *0.2*0.4750
   ≈ 0.1 + 0.1500 - 0.1649 + 0.0950 ≈ 0.1801

1.a.ii - Computed errors (betas) at each network node during the backward pass (9 marks)

1. Output node:
   βout = target - output = 1 - 0.1801 = 0.8199
2. Sigmoid 3:
   β3 = βout * w7 * y3 * (1 - y3)
   = 0.8199 * 0.2 * 0.4750 * (1 - 0.4750) ≈ 0.0409
3. Sigmoid 2:
   β2 = βout * w8 * y2 * (1 - y2)
   = 0.8199 * (-0.3) * 0.5498 * (1 - 0.5498) ≈ -0.0609
4. Sigmoid 1:
   β1 = βout * w9 * y1 * (1 - y1)
   = 0.8199 * 0.3 * 0.5000 * (1 - 0.5000) ≈ 0.0615

1.a.iii – Computed weight changes (deltas) and weight updates (9 marks)

1. Δw1 = η * β1 * x1 = 0.2 * 0.0615 * 0 = 0.0000
   w1(new) = -0.1 + 0.0000 = -0.1000
2. Δw2 = η * βout * x1 = 0.2 * 0.8199 * 0 = 0.0000
   w2(new) = -0.3 + 0.0000 = -0.3000
3. Δw3 = η * β2 * x1 = 0.2 * (-0.0609) * 0 = 0.0000
   w3(new) = -0.2 + 0.0000 = -0.2000
4. Δw4 = η * β2 * x2 = 0.2 * (-0.0609) * 1 ≈ -0.0122
   w4(new) = 0.2 + (-0.0122) ≈ 0.1878
5. Δw5 = η * βout * x2 = 0.2 * 0.8199 * 1 ≈ 0.1640
   w5(new) = 0.1 + 0.1640 = 0.2640
6. Δw6 = η * β3 * x2 = 0.2 * 0.0409 * 1 ≈ 0.0082
   w6(new) = -0.1 + 0.0082 = -0.0918
7. Δw7 = η * βout * y3 = 0.2 * 0.8199 * 0.4750 ≈ 0.0779
   w7(new) = 0.2 + 0.0779 = 0.2779
8. Δw8 = η * βout * y2 = 0.2 * 0.8199 * 0.5498 ≈ 0.0902
   w8(new) = -0.3 + 0.0902 = -0.2098
9. Δw9 = η * βout * y1 = 0.2 * 0.8199 * 0.5000 ≈ 0.0820
   w9(new) = 0.3 + 0.0820 = 0.3820
10. Δw10 = η * β3 * x1 = 0.2 * 0.0409 * 0 = 0.0000
    w10(new) = 0 + 0.0000 = 0.0000
11. Δw11 = η * β1 * x2 = 0.2 * 0.0615 * 1 ≈ 0.0123
    w11(new) = 0 + 0.0123 = 0.0123

Weight Changes:
w1: 0.0000, w2: 0.0000, w3: 0.0000, w4: -0.0122, w5: 0.1640, w6: 0.0082, w7: 0.0779, w8: 0.0902, w9: 0.0820, w10: 0.0000, w11: 0.0123

Updated Weights After First Input:
w1: -0.1000, w2: -0.3000, w3: -0.2000, w4: 0.1878, w5: 0.2640, w6: -0.0918, w7: 0.2779, w8: -0.2098, w9: 0.3820, w10: 0.0000, w11: 0.0123

1.b – Calculations to process the second training example (x1 = 1, x2 = 0, target = 1)

1.b.i – Computed outputs for each network node during the forward pass (4 marks)

1. Sigmoid 1 (y1):
net1 = w1 * x1 + w11 * x2 = -0.1000 * 1 + 0.0123 * 0 = -0.1000
y1 = 1 / (1 + e^(0.1000)) ≈ 0.4750

2. Sigmoid 2 (y2):
net2 = w3 * x1 + w4 * x2 = -0.2000 * 1 + 0.1878 * 0 = -0.2000
y2 = 1 / (1 + e^(0.2000)) ≈ 0.4502

3. Sigmoid 3 (y3):
net3 = w6 * x2 + w10 * x1 = -0.0918 * 0 + 0.0000 * 1 = 0.0000
y3 = 1 / (1 + e^(-0.0000)) = 0.5000

4. Sum Node (output):
output = w2x1 + w5x2 + w9y1 + w8y2 + w7y3
= -0.3000 * 1 + 0.2640 * 0 + 0.3820 * 0.4750 + (-0.2098) * 0.4502 + 0.2779 * 0.5000
≈ -0.3000 + 0 + 0.1815 - 0.0945 + 0.1390 ≈ -0.0740

1.b.ii - Computed errors (betas) at each network node during the backward pass (9 marks)

1. Output node:
βout = target - output = 1 - (-0.0740) = 1.0740

2. Sigmoid 3:
β3 = βout * w7 * y3 * (1 - y3)
= 1.0740 * 0.2779 * 0.5000 * (1 - 0.5000) ≈ 0.0746

3. Sigmoid 2:
β2 = βout * w8 * y2 * (1 - y2)
= 1.0740 * (-0.2098) * 0.4502 * (1 - 0.4502) ≈ -0.0558

4. Sigmoid 1:
β1 = βout * w9 * y1 * (1 - y1)
= 1.0740 * 0.3820 * 0.4750 * (1 - 0.4750) ≈ 0.1023

1.b.iii – Computed weight changes (deltas) and weight updates (9 marks)

1. $\Delta w1 = \eta * \beta1 * x1 = 0.2 * 0.1023 * 1 \approx 0.0205$
$w1(new) = -0.1000 + 0.0205 = -0.0795$

2. $\Delta w2 = \eta * \beta out * x1 = 0.2 * 1.0740 * 1 \approx 0.2148$
$w2(new) = -0.3000 + 0.2148 = -0.0852$

3. $\Delta w3 = \eta * \beta2 * x1 = 0.2 * (-0.0558) * 1 \approx -0.0112$
$w3(new) = -0.2000 + (-0.0112) = -0.2112$

4. $\Delta w4 = \eta * \beta2 * x2 = 0.2 * (-0.0570) * 0 = 0.0000$
$w4(new) = 0.1878 + 0.0000 = 0.1878$

5. $\Delta w5 = \eta * \beta out * x2 = 0.2 * 1.0740 * 0 = 0.0000$
$w5(new) = 0.2640 + 0.0000 = 0.2640$

6. $\Delta w6 = \eta * \beta3 * x2 = 0.2 * 0.0745 * 0 = 0.0000$
$w6(new) = -0.0918 + 0.0000 = -0.0918$

7. $\Delta w7 = \eta * \beta out * y3 = 0.2 * 1.0740 * 0.5000 \approx 0.1074$
$w7(new) = 0.2779 + 0.1074 = 0.3853$

8. $\Delta w8 = \eta * \beta out * y2 = 0.2 * 1.0740 * 0.4502 \approx 0.0967$
$w8(new) = -0.2098 + 0.0967 = -0.1131$

9. $\Delta w9 = \eta * \beta out * y1 = 0.2 * 1.0740 * 0.4750 \approx 0.1020$
$w9(new) = 0.3820 + 0.1020 = 0.4840$

10. $\Delta w10 = \eta * \beta3 * x1 = 0.2 * 0.0746 * 1 \approx 0.0149$
$w10(new) = 0.0000 + 0.0149 = 0.0149$

11. $\Delta w11 = \eta * \beta1 * x2 = 0.2 * 0.0927 * 0 = 0.0000$
$w11(new) = 0.0123 + 0.0000 = 0.0123$

Updated Weights After Second Input:
w1: -0.0795, w2: -0.0852, w3: -0.2112, w4: 0.1878, w5: 0.2640, w6: -0.0918, w7: 0.3853, w8: -0.1131, w9: 0.4840, w10: 0.0149, w11: 0.0123