Shubham Arya 1001650536

CSE 4309 – 001 Machine Learning
Assignment 4

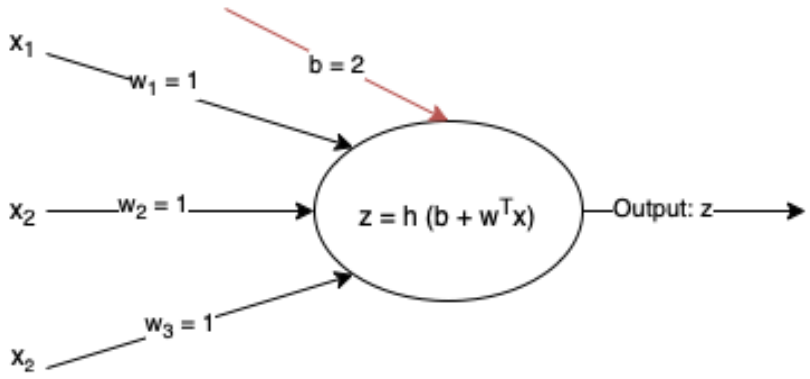
Task – 1

- Training and testing on pendigits dataset, with 2 layers, 10 training rounds.
  - classification accuracy=0.8659
- Training and testing on pendigits dataset, with 3 layers, 20 units per hidden layer, 20 training rounds.
  - classification accuracy=0.9188

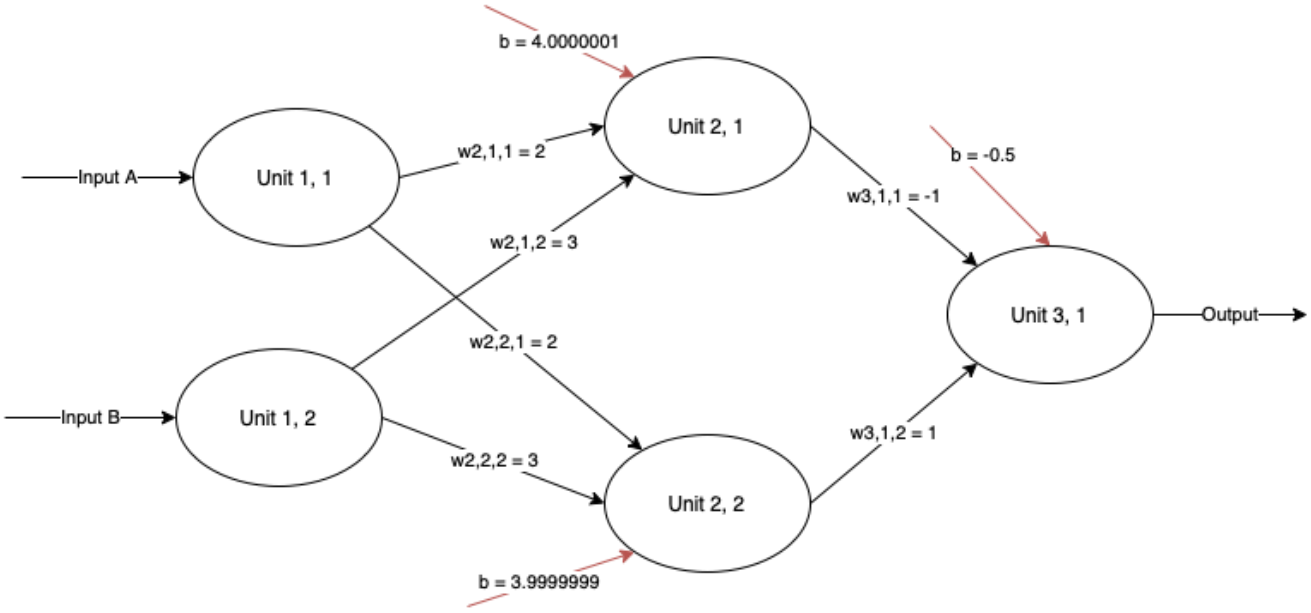
Task – 1b

The best test accuracy I received for pendigits dataset was for 3 layers, 30 units_per_layer and 70 rounds with a classification accuracy of 0.9280. These values took a long time to print so I am not sure if there were values over this that could give a better accuracy.

Task – 2



Task – 3



If 2A + 3B = 4,

Output of Unit 2,1 = h (2A + 3B – 4.00001) = h (4 – 4.00001) = 0
Output of Unit 2,2 = h (2A + 3B – 3.99999) = h (4 – 3.99999) = 1
Output of Unit 3,1 = h (1(1) - 1(0) – 0.5) = h (0.5) = **1**

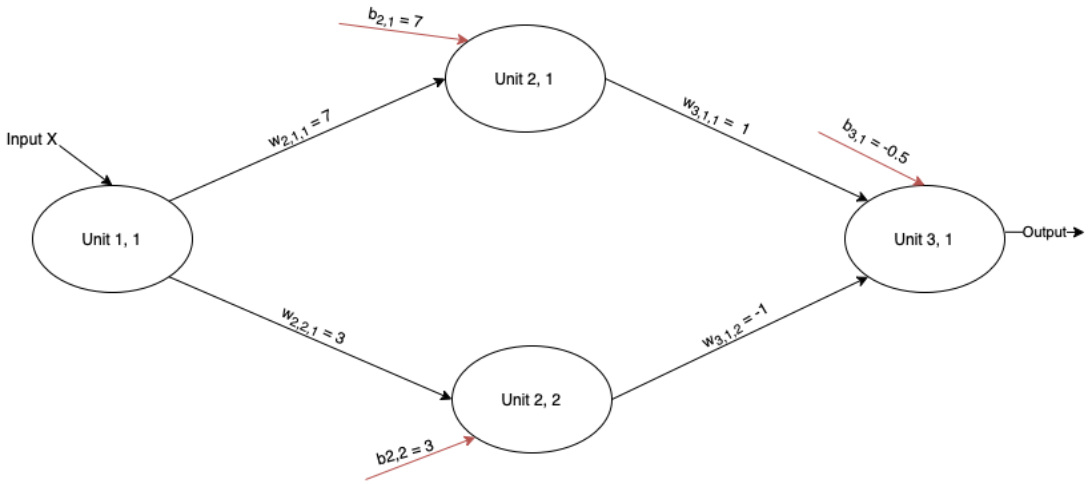
If 2A + 3B = 3,

Output of Unit 2,1 = h (2A + 3B – 4.00001) = h (3 – 4.00001) = 0
Output of Unit 2,2 = h (2A + 3B – 3.99999) = h (3 – 3.99999) = 0
Output of Unit 3,1 = h (1(0) - 1(0) – 0.5) = h (-0.5) = **0**

If 2A + 3B = 5,

Output of Unit 2,1 = h (2A + 3B – 4.00001) = h (5 – 4.00001) = 1
Output of Unit 2,2 = h (2A + 3B – 3.99999) = h (5 – 3.99999) = 1
Output of Unit 3,1 = h (1(1) - 1(1) – 0.5) = h (-0.5) = **0**

Task – 4



If X < 3, let's say X = 2

Output of Unit 2,1 = h (X – 7) = h (-5) = 0
Output of Unit 2,2 = h (X – 3) = h (-1) = 0
Output of Unit 3,1 = h (1(0) - 1(0) – 0.5) = h (-0.5) = **0**

If X > 7, let's say X = 8

Output of Unit 2,1 = h (X – 7) = h (1) = 1
Output of Unit 2,2 = h (X – 3) = h (5) = 1
Output of Unit 3,1 = h (1(1) - 1(1) – 0.5) = h (-0.5) = **0**

If X > 3 & X < 7, let's say X = 5

Output of Unit 2,1 = h (X – 7) = h (-2) = 0
Output of Unit 2,2 = h (X – 3) = h (2) = 1
Output of Unit 3,1 = h (1(1) - 1(0) – 0.5) = h (0.5) = **1**

Task – 5

If all weights are initialized to 0, it will lead to all neurons learning the same features during training. In such a case, the hidden layers will have identical influence on the cost which leads to identical gradient. In such a case, if the neurons have not learned a little bit of varying features, it could lead to a decrease in the classification accuracy.